

From: Sitton, Jehanne Souaya
Sent: Monday, July 10, 2006 4:38 PM
To: STIC-Biotech/ChemLib
Subject: sequence search request for 10/723,518

please perform the following sequence search:

please search positions 33601-34000 of SEQ ID NO: 1 where position 33670 is a G, in all nucleic acid databases. please include the first 50 hits and alignments or each.

thanks,
Jehanne Sitton
Primary Examiner
Art Unit 1634
Rem-2D81 (office)
Rem-2C70 (mailbox)
571-272-0752

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 17, 2006, 16:57:29 ; Search time 9797 Seconds

(Without alignments)
2610.898 Million cell updates/sec

Title: SEQ1-33670G

Perfect score: 399.2
Sequence: 1 aagaatgctctctctctata.....taactcagatcagaacgag 400

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 150 summaries

```

GenEmbl:*
1: gb_env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_ro:*
8: gb_ro:*
9: gb_un:*
10: gb_vl:*
11: gb_ov:*
12: gb_ptg:*
13: gb_in:*
14: gb_om:*
15: gb_ba:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Query Match	Length	ID	Description
1	397.6	99.6	33745	5	AY675320 Homo sapi
2	397.6	99.6	148386	5	AC104391 Homo sapi
3	397.6	99.6	193100	12	AC016729 Homo sapi
4	397.6	99.6	193363	2	CS200151 Sequence
5	397.6	99.6	193364	5	AC087350 Homo sapi
6	379.6	95.1	156441	5	AP004074 Homo sapi
7	62.2	15.6	211875	12	AC152315 Bos tauru
8	60.8	15.2	256751	6	AC122925 Mus muscu
9	60.6	15.2	120558	11	CR450689 Zebrafish
10	59.6	14.9	152637	12	AC151196 Mus muscu
11	59.6	14.9	181457	6	AC113480 Mus muscu
12	59	14.8	64239	12	AC130372 Homo sapi
13	58.8	14.7	104485	5	AC069439 Homo sapi
14	58.6	14.7	92354	5	AC106720 Homo sapi
15	58.6	14.7	147328	12	AC027755 Homo sapi
16	58.6	14.7	148851	12	AC016227 Homo sapi
17	58.6	14.7	303189	12	AC108728 Homo sapi
18	57.2	14.3	175262	5	AC021491 Homo sapi

19	56.6	14.2	149879	12	AC026572 Homo sapi
20	56.6	14.2	169897	11	AL929213 Zebrafish
21	56.6	14.2	211076	12	AC012214 Homo sapi
22	56.2	14.1	112727	6	AC098683 Mus muscu
23	56.2	14.1	118646	5	AC060796 Homo sapi
24	56.2	14.1	235471	6	AC156029 Mus muscu
25	56.2	14.1	236644	12	AC096353 Rattus no
26	56.2	14.1	239980	12	AC133731 Rattus no
27	56	14.0	319	6	AL585780 Taniarciu
28	56	14.0	171989	5	BS000611 Pan trogl
29	56	14.0	179557	5	AC151479 Pan trogl
30	56	14.0	224014	5	BS000534 Mus muscu
31	56	14.0	232304	5	AC146193 Pan trogl
32	55.8	14.0	60810	5	AC025227 Homo sapi
33	55.8	14.0	110139	11	CR749745 Zebrafish
34	55.8	14.0	212207	5	AC093853 Homo sapi
35	55.6	13.9	176108	5	AC026118 Homo sapi
36	55.6	13.9	184182	5	AC093515 Homo sapi
37	55.2	13.8	668	7	AL593617 Meleagris
38	55.2	13.8	668	7	AL593618 Meleagris
39	55.2	13.8	61255	12	AC093909 Mus muscu
40	55.2	13.8	151358	12	AL590437 Homo sapi
41	55.2	13.8	190080	5	AC022463 Homo sapi
42	55.2	13.8	197131	6	AC136742 Mus muscu
43	55.2	13.8	235887	12	AC106092 Rattus no
44	55	13.8	423	4	AF158547 Pinus tae
45	55	13.8	645	4	AF286606 Pinus tae
46	55	13.8	84422	12	BX088574_4 Continuation (5 of
47	55	13.8	107344	12	AC127799_3 Continuation (4 of
48	55	13.8	110000	12	BX088574_0 Continuation (3 of
49	55	13.8	160358	12	BX133034 Danio rer
50	55	13.8	236227	12	AC133707 Rattus no
51	54.8	13.7	75594	5	AP003093 Homo sapi
52	54.8	13.7	110000	12	AC128221 Rattus no
53	54.8	13.7	122331	5	AL356072 Human DNA
54	54.8	13.7	141529	12	AC025330 Homo sapi
55	54.8	13.7	153605	12	AC019360 Homo sapi
56	54.8	13.7	156481	11	CR376730 Zebrafish
57	54.8	13.7	157708	5	AP005357 Homo sapi
58	54.8	13.7	187098	6	AC166486 Mus muscu
59	54.8	13.7	191692	12	AC023834 Mus muscu
60	54.8	13.7	195959	12	AC034143 Homo sapi
61	54.8	13.7	196490	6	AC124766 Mus muscu
62	54.8	13.7	198254	12	AC016049 Homo sapi
63	54.8	13.7	225253	12	AC098950 Rattus no
64	54.8	13.7	226724	6	AC125199 Mus muscu
65	54.8	13.7	266823	12	AC132688 Rattus no
66	54.8	13.7	271848	12	BX470091 Danio rer
67	54.6	13.7	110000	12	AL928849_2 Continuation (3 of
68	54.6	13.7	110000	12	AL928849_3 Continuation (4 of
69	54.6	13.7	110159	6	AC068178 Homo sapi
70	54.6	13.7	187999	6	AL672268 Mouse DNA
71	54.6	13.7	196315	5	AC160873 Colobus g
72	54.6	13.7	211372	12	AC153428 Bos tauru
73	54.6	13.7	223309	6	AC079446 Mus muscu
74	54.6	13.7	241975	11	AL928668 Zebrafish
75	54.4	13.6	124524	5	AC117411 Homo sapi
76	54.4	13.6	149432	12	AC079309 Homo sapi
77	54.4	13.6	175775	5	AC006947 Homo sapi
78	54.4	13.6	175775	12	CR812943 Danio rer
79	54.4	13.6	237236	12	AC111444 Rattus no
80	54.4	13.6	237236	12	AL591666 Danio rer
81	54.2	13.6	162692	5	AC092958 Homo sapi
82	54.2	13.6	166491	12	AC021861 Homo sapi
83	54.2	13.6	221775	12	AC092992 Homo sapi
84	54.2	13.6	221794	12	AC020727 Homo sapi
85	54.2	13.6	228685	12	AC105531 Rattus no
86	54.2	13.6	230223	12	AC125554 Rattus no
87	54.2	13.6	232191	12	AC174022 Bos tauru
88	54.2	13.6	244221	12	AC135828 Rattus no
89	54.2	13.6	266833	12	AC172676 Bos tauru
90	54.2	13.6	301262	12	AC171554 Bos tauru
91	54	13.5	402	7	AB144168 Homo sapi

C	92	54	13.5	78693	5	AC114734	AC114734 Homo sapi
	93	54	13.5	109340	12	CT573349	CT573349 Danio rer
	94	54	13.5	140079	5	AC147288	AC147288 Pan trogl
	95	54	13.5	159649	11	BX927349	BX927349 Zebrafish
	96	54	13.5	161173	12	AC149115	AC149115 Papio anu
	97	54	13.5	165603	12	CT027765	CT027765 Danio rer
	98	54	13.5	172015	12	AC149194	AC149194 Papio anu
C	99	54	13.5	173071	11	CR925789	CR925789 Zebrafish
	100	53.8	13.5	372	7	AB147472	AB147472 Homo sapi
	101	53.8	13.5	45687	5	AC134730	AC134730 Homo sapi
	102	53.8	13.5	54073	12	AC133458	AC133458 Homo sapi
	103	53.8	13.5	114411	5	AC010599	AC010599 Homo sapi
	104	53.8	13.5	152555	5	AL354830	AL354830 Human DNA
	105	53.8	13.5	168524	12	AC024465	AC024465 Homo sapi
	106	53.8	13.5	188587	6	AC124464	AC124464 Mus muscu
C	107	53.8	13.5	200840	5	AL158072	AL158072 Human DNA
	108	53.8	13.5	222879	12	CR931761	CR931761 Danio rer
	109	53.8	13.5	225611	6	AC131724	AC131724 Mus muscu
	110	53.8	13.5	239363	12	AC121057	AC121057 Rattus no
	111	53.6	13.4	134841	5	AC018840	AC018840 Homo sapi
	112	53.6	13.4	145031	5	BS000601	BS000601 Pan trogl
	113	53.6	13.4	161872	12	CT027724	CT027724 Danio rer
	114	53.6	13.4	164376	12	AC010823	AC010823 Homo sapi
	115	53.6	13.4	168812	5	AC011327	AC011327 Homo sapi
	116	53.6	13.4	170577	5	AC149095	AC149095 Pan trogl
	117	53.6	13.4	172077	5	AC023514	AC023514 Homo sapi
	118	53.6	13.4	172765	5	AC146209	AC146209 Pan trogl
	119	53.6	13.4	173719	11	BX005313	BX005313 Zebrafish
	120	53.6	13.4	180299	5	AC018493	AC018493 Homo sapi
	121	53.6	13.4	181900	5	AC066606	AC066606 Homo sapi
	122	53.6	13.4	181989	5	BS000663	BS000663 Pan trogl
	123	53.6	13.4	188379	5	AC026200	AC026200 Homo sapi
	124	53.6	13.4	190170	12	AC149721	AC149721 Bos tauru
	125	53.6	13.4	202904	12	AP002077	AP002077 Homo sapi
	126	53.6	13.4	223910	6	AC159712	AC159712 Mus muscu
	127	53.4	13.4	119387	5	AC108128	AC108128 Homo sapi
	128	53.4	13.4	154765	5	AC114649	AC114649 Mus muscu
	129	53.4	13.4	154814	5	AC093533	AC093533 Homo sapi
	130	53.4	13.4	163604	14	AC092727	AC092727 Bos tauru
	131	53.4	13.4	170361	6	AC110554	AC110554 Mus muscu
	132	53.4	13.4	177570	11	CR354422	CR354422 Zebrafish
	133	53.4	13.4	187423	12	AC034175	AC034175 Homo sapi
	134	53.4	13.4	187487	6	AC159969	AC159969 Mus muscu
	135	53.4	13.4	195433	6	AC112925	AC112925 Mus muscu
	136	53.4	13.4	213461	12	AC182460	AC182460 Gallus ga
	137	53.4	13.4	217606	6	AC113176	AC113176 Mus muscu
	138	53.4	13.4	237754	12	AC117125	AC117125 Rattus no
	139	53.4	13.4	239559	12	AC095531	AC095531 Rattus no
	140	53.4	13.4	249062	12	AC160319	AC160319 Bos tauru
	141	53.4	13.4	250870	12	AC177806	AC177806 Gallus ga
	142	53.4	13.4	283738	12	AC178989	AC178989 Strongylo
	143	53.2	13.3	388	7	H8168XC3	Z67781 H. sapiens D
	144	53.2	13.3	113252	12	CT027769	Danio rer
	145	53.2	13.3	116250	6	AC153619	AC153619 Mus muscu
	146	53.2	13.3	140087	12	CR931815	Danio rer
	147	53.2	13.3	149596	12	AC009718	Mus muscu
	148	53.2	13.3	154123	5	AC138058	AC138058 Homo sapi
	149	53.2	13.3	155023	11	CR354440	CR354440 Zebrafish
	150	53.2	13.3	158636	12	CR450843	CR450843 Danio rer

ALIGNMENTS

RESULT 1
 AY675320/c 32745 bp DNA linear PRI 17-JUL-2004
 LOCUS Homo sapiens RAD21 homolog (S. pombe) (RAD21) gene, complete cds.
 DEFINITION AY675320
 ACCESSION AY675320.1 GI:50234988
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 32745) Livingston,R.J., Rieder,M.J., Chung,M.-W., Ritchie,T.K., Olson,A.N., Nguyen,C.P., Nguyen,D.A., Poel,C.V., Chambers,S.W., Schackwitz,W.S., Sherwood,J.K., Sherwood,A.M., Leitnauer,B.J. and Nickerson,D.A.
TITLE	Direct Submission
JOURNAL	Submitted (01-JUL-2004) Genome Sciences, University of Washington, 1705 NE Pacific, Seattle, WA 98195, USA
COMMENT	To cite this work please use: NIHES-SNPs, Environmental Genome Project, NIHES ES15478, Department of Genome Sciences, Seattle, WA (URL: http://sepp.gs.washington.edu).
FEATURES	Location/Qualifiers
source	1..32745 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"
repeat_region	1..1278 /rpt_family="U1" /rpt_type="dispersed"
variation	99 /frequency="0.02" /replace="g"
variation	120 /frequency="0.18" /replace="c"
variation	688 /frequency="0.03" /replace="c"
variation	846 /frequency="0.24" /replace="a"
variation	1016 /frequency="0.01" /replace="c"
variation	1062 /frequency="0.06" /replace="a"
variation	1179 /frequency="0.01" /replace="g"
variation	1248 /frequency="0.03" /replace="c"
variation	1338 /frequency="0.24" /replace="a"
variation	1350 /frequency="0.01" /replace="g"
variation	1422 /frequency="0.06" /replace="c"
variation	1464 /frequency="0.02" /replace="c"
variation	1815 /frequency="0.01" /replace="c"
variation	1817 /frequency="0.15" /replace="g"
variation	1946 /frequency="0.02" /replace="g"
variation	2243 /frequency="0.02" /replace="c"
variation	2262 /frequency="0.02" /replace="c"
variation	2268 /frequency="0.02" /replace="c"
variation	2276 /frequency="0.02" /replace="c"
variation	2282 /frequency="0.02" /replace="c"
variation	2288 /frequency="0.02" /replace="c"
variation	2294 /frequency="0.02" /replace="c"
variation	2300 /frequency="0.02" /replace="c"
variation	2306 /frequency="0.02" /replace="c"
variation	2312 /frequency="0.02" /replace="c"
variation	2318 /frequency="0.02" /replace="c"
variation	2324 /frequency="0.02" /replace="c"
variation	2330 /frequency="0.02" /replace="c"
variation	2336 /frequency="0.02" /replace="c"
variation	2342 /frequency="0.02" /replace="c"
variation	2348 /frequency="0.02" /replace="c"
variation	2354 /frequency="0.02" /replace="c"
variation	2360 /frequency="0.02" /replace="c"
variation	2366 /frequency="0.02" /replace="c"
variation	2372 /frequency="0.02" /replace="c"
variation	2378 /frequency="0.02" /replace="c"
variation	2384 /frequency="0.02" /replace="c"
variation	2390 /frequency="0.02" /replace="c"
variation	2396 /frequency="0.02" /replace="c"
variation	2402 /frequency="0.02" /replace="c"
variation	2408 /frequency="0.02" /replace="c"
variation	2414 /frequency="0.02" /replace="c"
variation	2420 /frequency="0.02" /replace="c"
variation	2426 /frequency="0.02" /replace="c"
variation	2432 /frequency="0.02" /replace="c"
variation	2438 /frequency="0.02" /replace="c"
variation	2444 /frequency="0.02" /replace="c"
variation	2450 /frequency="0.02" /replace="c"
variation	2456 /frequency="0.02" /replace="c"
variation	2462 /frequency="0.02" /replace="c"
variation	2468 /frequency="0.02" /replace="c"
variation	2474 /frequency="0.02" /replace="c"
variation	2480 /frequency="0.02" /replace="c"
variation	2486 /frequency="0.02" /replace="c"
variation	2492 /frequency="0.02" /replace="c"
variation	2498 /frequency="0.02" /replace="c"
variation	2504 /frequency="0.02" /replace="c"
variation	2510 /frequency="0.02" /replace="c"
variation	2516 /frequency="0.02" /replace="c"
variation	2522 /frequency="0.02" /replace="c"
variation	2528 /frequency="0.02" /replace="c"
variation	2534 /frequency="0.02" /replace="c"
variation	2540 /frequency="0.02" /replace="c"
variation	2546 /frequency="0.02" /replace="c"
variation	2552 /frequency="0.02" /replace="c"
variation	2558 /frequency="0.02" /replace="c"
variation	2564 /frequency="0.02" /replace="c"
variation	2570 /frequency="0.02" /replace="c"
variation	2576 /frequency="0.02" /replace="c"
variation	2582 /frequency="0.02" /replace="c"
variation	2588 /frequency="0.02" /replace="c"
variation	2594 /frequency="0.02" /replace="c"
variation	2600 /frequency="0.02" /replace="c"
variation	2606 /frequency="0.02" /replace="c"
variation	2612 /frequency="0.02" /replace="c"
variation	2618 /frequency="0.02" /replace="c"
variation	2624 /frequency="0.02" /replace="c"
variation	2630 /frequency="0.02" /replace="c"
variation	2636 /frequency="0.02" /replace="c"
variation	2642 /frequency="0.02" /replace="c"
variation	2648 /frequency="0.02" /replace="c"
variation	2654 /frequency="0.02" /replace="c"
variation	2660 /frequency="0.02" /replace="c"
variation	2666 /frequency="0.02" /replace="c"
variation	2672 /frequency="0.02" /replace="c"
variation	2678 /frequency="0.02" /replace="c"
variation	2684 /frequency="0.02" /replace="c"
variation	2690 /frequency="0.02" /replace="c"
variation	2696 /frequency="0.02" /replace="c"
variation	2702 /frequency="0.02" /replace="c"
variation	2708 /frequency="0.02" /replace="c"
variation	2714 /frequency="0.02" /replace="c"
variation	2720 /frequency="0.02" /replace="c"
variation	2726 /frequency="0.02" /replace="c"
variation	2732 /frequency="0.02" /replace="c"
variation	2738 /frequency="0.02" /replace="c"
variation	2744 /frequency="0.02" /replace="c"
variation	2750 /frequency="0.02" /replace="c"
variation	2756 /frequency="0.02" /replace="c"
variation	2762 /frequency="0.02" /replace="c"
variation	2768 /frequency="0.02" /replace="c"
variation	2774 /frequency="0.02" /replace="c"
variation	2780 /frequency="0.02" /replace="c"
variation	2786 /frequency="0.02" /replace="c"
variation	2792 /frequency="0.02" /replace="c"
variation	2798 /frequency="0.02" /replace="c"
variation	2804 /frequency="0.02" /replace="c"
variation	2810 /frequency="0.02" /replace="c"
variation	2816 /frequency="0.02" /replace="c"
variation	2822 /frequency="0.02" /replace="c"
variation	2828 /frequency="0.02" /replace="c"
variation	2834 /frequency="0.02" /replace="c"
variation	2840 /frequency="0.02" /replace="c"
variation	2846 /frequency="0.02" /replace="c"
variation	2852 /frequency="0.02" /replace="c"
variation	2858 /frequency="0.02" /replace="c"
variation	2864 /frequency="0.02" /replace="c"
variation	2870 /frequency="0.02" /replace="c"
variation	2876 /frequency="0.02" /replace="c"
variation	2882 /frequency="0.02" /replace="c"
variation	2888 /frequency="0.02" /replace="c"
variation	2894 /frequency="0.02" /replace="c"
variation	2900 /frequency="0.02" /replace="c"
variation	2906 /frequency="0.02" /replace="c"
variation	2912 /frequency="0.02" /replace="c"
variation	2918 /frequency="0.02" /replace="c"
variation	2924 /frequency="0.02" /replace="c"
variation	2930 /frequency="0.02" /replace="c"
variation	2936 /frequency="0.02" /replace="c"
variation	2942 /frequency="0.02" /replace="c"
variation	2948 /frequency="0.02" /replace="c"
variation	2954 /frequency="0.02" /replace="c"
variation	2960 /frequency="0.02" /replace="c"
variation	2966 /frequency="0.02" /replace="c"
variation	2972 /frequency="0.02" /replace="c"
variation	2978 /frequency="0.02" /replace="c"
variation	2984 /frequency="0.02" /replace="c"
variation	2990 /frequency="0.02" /replace="c"
variation	2996 /frequency="0.02" /replace="c"
variation	3002 /frequency="0.02" /replace="c"
variation	3008 /frequency="0.02" /replace="c"
variation	3014 /frequency="0.02" /replace="c"
variation	3020 /frequency="0.02" /replace="c"
variation	3026 /frequency="0.02" /replace="c"
variation	3032 /frequency="0.02" /replace="c"
variation	3038 /frequency="0.02" /replace="c"
variation	3044 /frequency="0.02" /replace="c"
variation	3050 /frequency="0.02" /replace="c"
variation	3056 /frequency="0.02" /replace="c"
variation	3062 /frequency="0.02" /replace="c"
variation	3068 /frequency="0.02" /replace="c"
variation	3074 /frequency="0.02" /replace="c"
variation	3080 /frequency="0.02" /replace="c"
variation	3086 /frequency="0.02" /replace="c"
variation	3092 /frequency="0.02" /replace="c"
variation	3098 /frequency="0.02" /replace="c"
variation	3104 /frequency="0.02" /replace="c"
variation	3110 /frequency="0.02" /replace="c"
variation	3116 /frequency="0.02" /replace="c"
variation	3122 /frequency="0.02" /replace="c"
variation	3128 /frequency="0.02" /replace="c"
variation	3134 /frequency="0.02" /replace="c"
variation	3140 /frequency="0.02" /replace="c"
variation	3146 /frequency="0.02" /replace="c"
variation	3152 /frequency="0.02" /replace="c"
variation	3158 /frequency="0.02" /replace="c"
variation	3164 /frequency="0.02" /replace="c"
variation	3170 /frequency="0.02" /replace="c"
variation	3176 /frequency="0.02" /replace="c"
variation	3182 /frequency="0.02" /replace="c"
variation	3188 /frequency="0.02" /replace="c"
variation	3194 /frequency="0.02" /replace="c"
variation	3200 /frequency="0.02" /replace="c"
variation	3206 /frequency="0.02" /replace="c"
variation	3212 /frequency="0.02" /replace="c"
variation	3218 /frequency="0.02" /replace="c"
variation	3224 /frequency="0.02" /replace="c"
variation	3230 /frequency="0.02" /replace="c"
variation	3236 /frequency="0.02" /replace="c"
variation	3242 /frequency="0.02" /replace="c"
variation	3248 /frequency="0.02" /replace="c"
variation	3254 /frequency="0.02" /replace="c"
variation	3260 /frequency="0.02" /replace="c"
variation	3266 /frequency="0.02" /replace="c"
variation	3272 /frequency="0.02" /replace="c"
variation	3278 /frequency="0.02" /replace="c"
variation	3284 /frequency="0.02" /replace="c"
variation	3290 /frequency="0.02" /replace="c"
variation	3296 /frequency="0.02" /replace="c"
variation	3302 /frequency="0.02" /replace="c"
variation	3308 /frequency="0.02" /replace="c"
variation	3314 /frequency="0.02" /replace="c"
variation	

Db	4027	ATGCAGTTTCAGTTTCCCTACCTATGAAAACTGGAAAGTTGAAACTGTGTTTACTCTAA	3968
0Y	361	CCTTGAAAAAGTCATATCCTTACTCAGATCAGAACGAG	400
Db	3967	CCTTGAAAAAGTCATATCCTTACTCAGATCAGAACGAG	3928
RESULT 2			
AC104391/c			
LOCUS	AC104391	148386 bp	DNA linear PRI 01-MAY-2002
DEFINITION	Homo sapiens chromosome 8, clone CTD-3071X10, complete sequence.		
ACCESSION	AC104391		
VERSION	AC104391.5	GI:20377041	
KEYWORDS	HTC.		
ORGANISM	Homo sapiens (human)		
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 148386)		
AUTHORS	Bitren,B., Linton,L., Nusbaum,C. and Lander,E.		
TITLE	Homo sapiens chromosome 8, clone CTD-3071X10 unpublished		
JOURNAL	2 (bases 1 to 148386)		
REFERENCE	1 (bases 1 to 148386)		
AUTHORS	Bitren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S., Barna,N., Baetien,V., Bogunlavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karataas,A., Kells,C., Lacroque,K., Lamasares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Major,J., Marguis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McNeeters,R., Meldrim,J., Menau,L., Mhova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Noddu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Roesti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Sudbramanian,A., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zaitoun,J., Zembek,L., Zimmer,A. and Zody,M.		
TITLE	Submitted (08-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
JOURNAL	3 (bases 1 to 148386)		
REFERENCE	1 (bases 1 to 148386)		
AUTHORS	Bitren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S., Barna,N., Baetien,V., Bloom,T., Bogunlavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karataas,A., Kells,C., Lacroque,K., Lamasares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Major,J., Marguis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McNeeters,R., Meldrim,J., Menau,L., Mhova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Noddu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Roesti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Sudbramanian,A., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zaitoun,J., Zembek,L., Zimmer,A. and Zody,M.		

```

TITLE      Direct Submission
JOURNAL    Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome
REFERENCE  Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS    4 (bases 1 to 148386)
           Bliren,B., Linton,L., Nudbaum,C., Lander,E., All,A., Allen,N.,
           Anderson,S., Barna,N., Bastien,Y., Bloom,T., Boguslavsky,L.,
           Boudgaler,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
           Chazaro,B., Chopel,Y., Colangelo,M., Collins,S., Collymore,A.,
           Cook,K., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
           Fero,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
           Ginde,S., Gordon,S., Goyette,M., Graham,L., Grand-Pierre,N.,
           Hagos,B., Horton,L., Hulme,W., Iller,I., Johnson,R., Jones,C.,
           Kamat,A., Karasch,A., Kells,C., Lacroque,K., Lamazares,R.,
           Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
           Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
           McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Mense,L.,
           Michova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
           Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
           Oliver,J., Peterson,K., Punthang,P., Pierre,N., Pollard,V.,
           Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
           Roman,J., Rossetti,M., Roy,A., Santos,R., Schauer,S., Schuppach,R.,
           Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
           Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
           Tophan,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
           Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
           Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

COMMENT    Direct Submission
           Submitted (01-MAY-2002) Whitehead Institute/MIT Center for Genome
           Research, 320 Charles Street, Cambridge, MA 02141, USA
           On May 1, 2002 this sequence version replaced gi:20128309.
           All repeats were identified using RepeatMasker:
           Smit, A.F.A. & Green, P. (1996-1997)
           http://ftp.genome.washington.edu/RM/RepeatMasker.html

           ----- Genome Center
           Center: Whitehead Institute/ MIT Center for Genome Research
           Center code: WIBR
           Web site: http://www-seq.wi.mit.edu
           Contact: sequence_submissions@genome.wi.mit.edu
           ----- Project Information
           Center project name: L22223
           Center clone name: 3071_X_10

           -----
           Location/Qualifiers
               1..148386
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
               /chromosome="8"
               /map="8"
               /clone="CTD-3071K10"
               /clone_id="CTD2 Human BAC"
               /complement(336..388)
               /rpt_family="L2"
               /complement(830..1376)
               /rpt_family="MLTK"
               /complement(1380..1619)
               /rpt_family="MIR"
               /complement(1911..1943)
               /rpt_family="MIR"
               /complement(2308..2316)
               /note="<30 qual SINGL region"
               /rpt_family="AluBx"
               /complement(3365..3815)
               /rpt_family="MLTK"
               4200..4239
               /note="single clone coverage"
               4418..4558
               /rpt_family="L2"
               5746..5797
               /rpt_family="GA-rich"
               5820..5975
               /rpt_family="L1MA8"

```

OY 1 AAGATGCTTTCTCTATATTTAAGAGTCTTGATAGCTTAAATCCTGGCGC 60
 Db 129948 AAGATGCTTTCTCTATATTTAAGAGTCTTGATAGCTTAAATCCTGGCGC 129885

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Project Information -----
Center project name: H_NH0367C15
----- Summary Statistics -----
Sequencing vector: M13; 76%
Sequencing vector: plasmid; 24%
Chemistry: Dye-primer ET; 76% of reads
Chemistry: Dye-terminator Bis Dye; 24% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 186568 bases at least Q40
Consensus quality: 189080 bases at least Q30
Consensus quality: 190451 bases at least Q20
Insert size: 199000; agarose-fp
Insert size: 192300; sum-of-contigs
Quality coverage: 4.33 in Q20 bases; agarose-fp
Quality coverage: 4.86 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is

```

```

* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1      1597: contig of 1597 bp in length
*      1598      1697: gap of unknown length
*      1698      8306: contig of 6609 bp in length
*      8307      8406: gap of unknown length
*      8407      15335: contig of 6929 bp in length
*      15336      15435: gap of unknown length
*      15436      22120: contig of 6685 bp in length
*      22121      22220: gap of unknown length
*      22221      30889: contig of 8669 bp in length
*      30890      30989: gap of unknown length
*      30990      39394: contig of 8405 bp in length
*      39395      39494: gap of unknown length
*      39495      56109: contig of 16615 bp in length
*      56110      56209: gap of unknown length
*      56210      99245: contig of 43036 bp in length
*      99246      99346: gap of unknown length
*      99346      193310: contig of 93755 bp in length.
*
* Location/Qualifiers

```

source	1..193100
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="8"
	/clone="RP11-367C15"
	1..1597
misc_feature	/note="assembly_name:Contig3"
gap	1598..1697
	/estimated_length=unknown
misc_feature	1698..8306
	/note="assembly_name:Contig4"
gap	8307..8406
	/estimated_length=unknown
misc_feature	8407..15335
	/note="assembly_name:Contig5"
gap	15336..15435
	/estimated_length=unknown
misc_feature	15436..22120
	/note="assembly_name:Contig6"
	clone_end:T7
gap	vector_side:right"
	22121..22220
	/estimated_length=unknown
misc_feature	22221..30889
	/note="assembly_name:Contig7"
gap	30890..30989
	/estimated_length=unknown
misc_feature	30990..39394
	/note="assembly_name:Contig8"
gap	39395..39494
	/estimated_length=unknown
misc_feature	39495..56109
	/note="assembly_name:Contig9"
gap	56110..56209
	/estimated_length=unknown
misc_feature	56210..99245
	/note="assembly_name:Contig10"
gap	99246..99345
	/estimated_length=unknown
misc_feature	99346..193100
	/note="assembly_name:Contig11"
	clone_end:SP6
	vector_side:left"

Query Match	99.6%	Score 397.6	DB 12	Length 193100
Best Local Similarity	99.2%	Pred. 2.8e-123		
Matches 397	Conservative	2	Mismatches 1	Indels 0
			Gaps	0

QY	1	AAGATGCTCTTTCCTATATATATTTACAGAGCTTGAATPAACCTPAAATATCCTGGGCG	60
Db	192520	AAGAAATGCTCTTTCCTATATATATTTACAGAGCTTGAATPAACCTPAAATATCCTGGGCG	192461
QY	61	TACACATGAGACTGSGTAGTACTGTTACCTCTGGAAGTAGTATGACAGGCAAGTAGG	120
Db	192460	TACACATGAAACTGSGTAGTACTGTTACCTCTGGAAGTAGGATATGACAGGCAAGTAGG	192401
QY	121	CCTACCTTTTGTACTGTTGATATATATATTTACATACACACACACACACACA	180
Db	192400	CCTACCTTTTGTACTGTTGATATATATATTTACATACACACACACACACACA	192341
QY	181	CAVATATTTTAACTGGGGACATAAATGTACACACCACTATTCAAAGTAAATCTACTCT	240
Db	192340	CACATATTTTAACTGGGGACATAAATGTACACCACTATTCAAAGTAAATCTACTACT	192281
QY	241	CAGAAGTAAAGACTATATCTGCTATTTGTCAATTAATACAGATTAAAGTCCCTAACCTT	300
Db	192280	CAGAAGTAAAGACTATATCTGCTATTTGTCAATTAATACAGATTAAAGTCCCTAACCTT	192221
QY	301	ATGCAAGTTCAGTTTCCCTATCTATGAAAACTGAAAGGTTGAAATCTGGTGTTACCTAA	360
Db	192220	ATGCAAGTTCAGTTTCCCTACTATGAAAACTGAAAGGTTGAAATCTGGTGTTACCTAA	192161
QY	361	CCTTGGAAAAAGCTCAATACCTATCTCAAGTACGAACGGAG	400
Db	192160	CCTTGGAAAAAGCTCAATACCTATCTCAAGTACGAACGGAG	192121

```

RESULT 4
CS200151          CS200151      193363 bp      DNA      linear      PAT 08-DEC-20
LOCUS
DEFINITION       Sequence 32 from Patent WO2005106044.
ACCESSION        CS200151
VERSION          CS200151.1  GI:83408556
KEYWORDS
SOURCE
ORGANISM         Homo sapiens (human)
                 Homo sapiens
                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
                 Homnidae; Homo.
REFERENCE
1
AUTHORS         Harris, C.
JOURNAL         Patent: WO 2005106044-A 32 10-NOV-2005;
                 Exogen Diagnostics, Inc. (US)
FEATURES
source
                 1..193363
                 location/Qualifiers
                 /organism="Homo sapiens"
                 /mol_type="unassigned DNA"
                 /db_xref="taxon:9606"
ORIGIN

```

	Query Match	99.6%	Score 397.6	DB 2	Length 193363
	Best Local Similarity	99.2%	Pred. 2.8e-123		
	Matches 397	Conservative	2	Mismatches 1	Indels 0; Gaps 0;
Qy	1 AAGAAATGCTCTTTCCTATATATTACGAATCTTGTAATAAATACCTGTGGCG	60			
Db	100161 AAGAATGCTTTCTCTATATATTACGAATCTTGTAATAAATACCTGTGGCG	100220			
Qy	61 TACACATGACCTGSGATAGTACTGTTACCTCTGGAAAGTAGATGCACGGCAGAGAG	120			
Db	100221 TACACATGAACTGGTGAATAGTACTGTTACCTCTGGAAAGTAGATGCACGGCAGAGAG	100280			
Qy	121 CCTACCTTTTGTACTGTTGATATATATATATATACACACACACACACACA	180			
Db	100281 CCTACCTTTTGTACTGTTGATATATATATATATACACACACACACACACA	100340			
Qy	181 CAVATTTTTAACCTGGGACATAAAGTACACACCCTATTCCAAAGTAAAACTACTATCT	240			
Db	100341 CACATTTTTTAACCTGGGACATAAAGTACACACCCTATTCCAAAGTAAAACTACTATCT	100400			
Qy	241 CAGAACTAAAGCATATATCTCTGATATGCAATTAACAGCAAGTAAAGCTCAACCTTT	300			

OY		100461	CAGAACTAAGACATTAATCCTGCTATTCGTGCATTACTAGCATGAAGTAACTGGTGTTTACCCTTA	360
D5		100461	ATGCAGTTCAAGTTTTCCCTACATCTAGTATGAAACCTGAAAGCTGAAATCTGGTGTTTACCCTTA	360
OY				
D5		100461	ATGCAGTTCAAGTTTTCCCTACATCTAGTATGAAACCTGAAAGCTGAAATCTGGTGTTTACCCTTA	100520
OY				
D5		100521	CCTTTGAAAACCGTCAATACCTTAACCTGATGATGCAAGGAG	100560
RESULT 5				
LOCUS	AC087350/c	193364 bp	DNA	linear PRI 21-AUG-2002
DEFINITION	Homo sapiens chromosome 8, clone RP11-367C15, complete sequence.			
ACCESSION	AC087350			
VERSION	AC087350.3	GI:22267854		
KEYWORDS	HTG.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 193364) Birren,B., Nusbaum,C. and Lande,E. Homo sapiens chromosome 8, clone RP11-367C15 Unpublished 2 (bases 1 to 193364) Birren,B., Linton,L., Nusbaum,C., Lande,E., Allen,N., Anderson,S., Barna,N., Baetien,V., Boguslavsky,L., Bonkhalter,B., Brown,A., Camrata,U., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyn,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heatord,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karacas,A., Labocque,K., Lamarez,R., Landers,T., Lehoczy,U., Levine,R., Liu,G., Maclean,C., MacDonald,P., Marquis,N., Matthews,C., McCarthy,M., Mcwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneis,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,F., O'Neil,D., Oliver,J., Peterson,K., PhunKhang,P., Pletre,N., Pollara,V., Raymond,C., Retta,R., Rinback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Souner,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Travers,N., Travis,N., Ttiglio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.G., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.			
TITLE	Direct Submission			
JOURNAL	Submitted (28-DEC-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA			
REFERENCE	3 (bases 1 to 193364) Birren,B., Nusbaum,C., Lande,E., Ali,A., Allen,N., Anderson,S., Barna,N., Baetien,V., Bloom,T., Boguslavsky,L., Bonkhalter,B., Camarata,U., Chang,U., Chazarro,K., Choepel,Y., Collymore,A., Cooke,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karakas,A., Kelle,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., MacDonald,P., Maylor,J., Matthews,C., McCarthy,M., Meldrim,J., Meneis,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,F., O'Neil,D., Oliver,J., Peterson,K., PhunKhang,P., Pletre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Tophan,K., Travers,N., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.			
AUTHORS	Direct Submission			
TITLE	Direct Submission			

JOURNAL	Submitted (16-MUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	4 (bases 1 to 193364)
AUTHORS	Barnett, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barta, N., Basset, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Comarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., DeRubeano, K., Dewar, K., Diaz, J. S., Dodge, S., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagob, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Menue, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Plunkharg, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Talamas, J., Tsafaye, S., Theodore, J., Tophan, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
TITLE	Direct Submissions
JOURNAL	Submitted (21-MUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Aug 16, 2002 this sequence version replaced gi:16905274. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center -----	
Center: Whitehead Institute/ MIT Center for Genome Research	
Center code: MIBR	
Web site: http://www-seq.wi.mit.edu	
Contact: sequence.submissions@genome.wi.mit.edu	
----- Project Information -----	
Center project name: L1698	
Center clone name: 367_C15	

FEATURES	Location/Qualifiers
SOURCE	1. 193364
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/ID_xref="taxon:9606"
	/chromosome="8"
	/map="8"
	/clone="RP11-367C15"
	/clone_lib="RPC1-11 Human Male BAC"
repeat_region	230..348
	/rpt_family="MIR"
repeat_region	349..651
	/rpt_family="AluSq"
repeat_region	652..730
	/rpt_family="MIR"
repeat_region	1721..2290
	/rpt_family="L1MB5"
repeat_region	3577..3711
	/rpt_family="MSTD"
repeat_region	3728..3786
	/rpt_family="PRAM"
repeat_region	3812..4016
	/rpt_family="MSTD"
repeat_region	complement(4135..4438)
	/rpt_family="AluY"
repeat_region	4531..4597
	/rpt_family="(TTCC)n"
repeat_region	5406..6437
	/rpt_family="L1MBC"
repeat_region	6438..6744
	/rpt_family="AluY"
repeat_region	6745..7328
	/rpt_family="L1MBC"
repeat_region	8016..8321
	/rpt_family="AluX"
repeat_region	complement(8338..8431)

OY		1	AAGAATGCTCCTTTTCTCATTAATTATAGAGTCCTGAATAAGCTAAATACCGTGAGCG	60
Dd	93203		AAGAATGCTCCTTTTCTCATTAATTATAGAGTCCTGAATAAGCTAAATACCGTGAGCG	9314
OY		61	TACAATGAGACTGTGTATACCTGGTTAACCCTGGAAAAGTAGATGCAGGCCAAGTAGG	120
Dd	93143	:	TACAATGAAAACTGTGTATACCTGGTTAACCCTGGAAAAGTAGATGCAGGCCAAGTAGG	9308
OY		121	CCTAACCTTTTTGTACTGTGTATATATATATATATATCAATCACACACACACACACACA	180
Dd	93083		CCTAACCTTTTTGTACTGTGTATATATATATATATATCAATCACACACACACACACACA	9302
OY		181	CAVATATTTTAACTCCGGGACATAAATGATACACCAACCTATTCAAAGTAAAACTACTATCT	240
Dd	93023		CACATATTTTAACTCCGGGACATAAATGATACACCAACCTATTCAAAGTAAAACTACTATCT	9286
OY		241	CAGAAGTAAAGACATTAATCCTGCTATTGTCAATTACTAGACAGTTAACTGCTTAACCTTT	300
Dd	92963		CAGAAGTAAAGACATTAATCCTGCTATTGTCAATTACTAGACAGTTAACTGCTTAACCTTT	9280
OY		301	ATGCAGTTCACTTCCCCTACTATATGAAAACTGAAAAGTTAAAACTGTGTATTACCTAA	360
Dd	92903		ATGCAGTTCACTTCCCCTACTATATGAAAACTGAAAAGTTAAAACTGTGTATTACCTAA	9284
OY		361	CCTTGGAAAAAGTCATATACCTTAATCTGATCAGATCAGAACCGAG	400
Dd	92843		CCTTGGAAAAAGTCATATACCTTAATCTGATCAGATCAGAACCGAG	92804
RESULT 6 AP004074 LOCUS			156441 bp DNA linear PRI 21-AUG-2001 sequence.	
DEFINITION	Homo sapiens genomic DNA, chromosome 8q23, clone:KB1266f9, complete			
ACCESSION	AP004074			
VERSION	AP004074.1 GI:15216353			
KEYWORDS	.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo. 1			
REFERENCE	Shimizu,N. and Asakawa,S. Homo sapiens DNA chromosome 8 SEQUENCE Published Only in Database (2001) 2 (bases 1 to 156441) Shimizu,N. and Asakawa,S. Direct Submission Submitted (20-AUG-2001) Nobuyoshi Shimizu, Keio University, School of Medicine, Molecular Biology / 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan (E-mail:nshimid@med.keio.ac.jp, Tel:81-3-3351-2370, Fax:81-3-3351-2370)			
JOURNAL TITLE				
FEATURES SOURCE				
source	1..156441 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="8" /map="Bq23" /clone="KB1266F9" /cell_line="PLEB 14-14" /clone_id="Keio BAC library"			
repeat_region	560..581 /inference="non-experimental evidence, no additional details recorded"			
repeat_region	/rpt_family="Al-rich" 2623..3033 /inference="non-experimental evidence, no additional details recorded"			
repeat_region	/rpt_family="MSTB" 3626..3688			

repeat_region	details recorded" /rpt_family="A)n" 5506..5537 /inference="non-experimental evidence, no additional details recorded" /rpt_family="CT-rich" 7579..7632 /inference="non-experimental evidence, no additional details recorded"	repeat_region	/rpt_family="Tigger1" 14872..15595 /inference="non-experimental evidence, no additional details recorded" /rpt_family="Tigger1" 15637..15922 /inference="non-experimental evidence, no additional details recorded"
repeat_region	details recorded" /rpt_family="AT-rich" complement(7956..8247) /inference="non-experimental evidence, no additional details recorded" /rpt_family="AluSx" 8697..8731 /inference="non-experimental evidence, no additional details recorded"	repeat_region	/rpt_family="L1MC5" 17431..17469 /inference="non-experimental evidence, no additional details recorded"
repeat_region	details recorded" /rpt_family="CAA)n" 9452..9560 /inference="non-experimental evidence, no additional details recorded"	repeat_region	/rpt_family="AT-rich" 19504..19526 /inference="non-experimental evidence, no additional details recorded"
repeat_region	details recorded" /rpt_family="L2" 10358..10385 /inference="non-experimental evidence, no additional details recorded"	repeat_region	/rpt_family="L1MC1" complement(22612..22919) /inference="non-experimental evidence, no additional details recorded"
repeat_region	details recorded" /rpt_family="AT-rich" 10421..10441 /inference="non-experimental evidence, no additional details recorded"	repeat_region	/rpt_family="AluTo" complement(23100..23442) /inference="non-experimental evidence, no additional details recorded"
repeat_region	details recorded" /rpt_family="AT-rich" complement(10443..10592) /inference="non-experimental evidence, no additional details recorded"	repeat_region	/rpt_family="L1MD1" 23442..24177 /inference="non-experimental evidence, no additional details recorded"
repeat_region	details recorded" /rpt_family="FLAM_C" 11163..11185 /inference="non-experimental evidence, no additional details recorded"	repeat_region	/rpt_family="L1MC/D" 24178..24474 /inference="non-experimental evidence, no additional details recorded"
repeat_region	details recorded" /rpt_family="A)n" 11849..11876 /inference="non-experimental evidence, no additional details recorded"	repeat_region	/rpt_family="AluY" 24475..24685 /inference="non-experimental evidence, no additional details recorded"
repeat_region	details recorded" /rpt_family="AT-rich" 12699..12861 /inference="non-experimental evidence, no additional details recorded"	repeat_region	/rpt_family="L1MC/D" 25120..25255 /inference="non-experimental evidence, no additional details recorded"
repeat_region	details recorded" /rpt_family="FAM" 13403..13423 /inference="non-experimental evidence, no additional details recorded"	repeat_region	/rpt_family="FLAM_C" complement(26608..26781) /inference="non-experimental evidence, no additional details recorded"
repeat_region	details recorded" /rpt_family="AT-rich" 13609..13719 /inference="non-experimental evidence, no additional details recorded"	repeat_region	/rpt_family="MIR" complement(26999..27171) /inference="non-experimental evidence, no additional details recorded"
repeat_region	details recorded" /rpt_family="MIR" 13731..13758 /inference="non-experimental evidence, no additional details recorded"	repeat_region	/rpt_family="MER2" 27399..27410 /inference="non-experimental evidence, no additional details recorded"
repeat_region	details recorded" /rpt_family="L1MC5" 14034..14260 /inference="non-experimental evidence, no additional details recorded"	repeat_region	/rpt_family="AT-rich" 27411..27435 /inference="non-experimental evidence, no additional details recorded"
repeat_region	details recorded" /rpt_family="Tigger1" complement(14263..14407) /inference="non-experimental evidence, no additional details recorded"	repeat_region	/rpt_family="CA)n" 29520..29564 /inference="non-experimental evidence, no additional details recorded"
repeat_region	details recorded" /rpt_family="FLAM_A" 14409..14877 /inference="non-experimental evidence, no additional details recorded"	repeat_region	/rpt_family="G-rich" 30299..32109 /inference="non-experimental evidence, no additional details recorded"
repeat_region	details recorded"	repeat_region	/rpt_family="L1MA4"

[illegible]

TITLE	REFERENCE	AUTHORS	JOURNAL
Unpublished	2 (bases 1 to 211875)	Worley, K.C.	Direct Submission
Submitted (01-NOV-2004)	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	3 (bases 1 to 211875)	Cow Genome Sequencing Consortium.
Direct Submission	Submitted (01-JUL-2005)	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	On Jun 26, 2005 this sequence version replaced gi:58038130. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence conflicts within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
Genome Center	Center: Baylor College of Medicine	Center code: BCM	Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu	Project Information	Center project name: FABZ	Center clone name: CH240-1K23
Summary Statistics	Assembly program: Atlas 3.0;		

Consensus quality: 198612 bases at least Q40
 Consensus quality: 201789 bases at least Q30
 Consensus quality: 204577 bases at least Q20
 Estimated insert size: 203572; sum-of-coverage estimation
 Quality coverage: 5x in Q20 bases; sum-of-coverage estimation

NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 NOTE: This is a 'working draft' sequence. It currently
 consists of 32 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 30291: contig of 30291 bp in length
 30292 30341: gap of 50 bp
 30342 37665: contig of 7324 bp in length
 37666 37715: gap of 50 bp
 37716 53826: contig of 16111 bp in length
 53827 53876: gap of 50 bp
 53877 56440: contig of 2564 bp in length
 56441 56490: gap of 50 bp
 56491 67817: contig of 11327 bp in length
 67818 68524: gap of 707 bp
 68525 70609: contig of 2085 bp in length
 70610 71044: gap of 435 bp
 71045 73476: contig of 2432 bp in length
 73477 73526: gap of 50 bp
 73527 92741: contig of 19215 bp in length
 92742 92791: gap of 50 bp
 92792 96778: contig of 3987 bp in length
 96779 96828: gap of 50 bp
 96829 105197: contig of 8369 bp in length
 105198 105297: gap of unknown length
 105298 106706: contig of 1409 bp in length
 106707 106756: gap of 50 bp
 106757 115759: contig of 9003 bp in length
 115760 115859: gap of unknown length
 115860 122530: contig of 6671 bp in length
 122531 122662: gap of 132 bp
 126722: contig of 4060 bp in length
 126723 126772: gap of 50 bp
 126773 131853: contig of 5081 bp in length
 131854 132386: gap of 533 bp
 132387 140048: contig of 7662 bp in length
 140049 140098: gap of 50 bp
 140099 148892: contig of 8794 bp in length
 148893 148950: gap of 58 bp
 148951 159941: contig of 10991 bp in length
 159942 159991: gap of 50 bp
 159992 167844: contig of 7853 bp in length
 167845 167944: gap of unknown length
 167945 177933: contig of 9989 bp in length
 177934 178226: gap of 293 bp
 178227 182916: contig of 4690 bp in length
 182917 182966: gap of 50 bp
 182967 185344: contig of 2378 bp in length
 185345 185394: gap of 50 bp
 185395 189477: contig of 4083 bp in length
 189478 189527: gap of 50 bp
 189528 192243: contig of 2716 bp in length
 192244 192293: gap of 50 bp
 192294 195792: contig of 3499 bp in length
 195793 195842: gap of 50 bp
 195843 204127: contig of 8285 bp in length
 204128 204227: gap of unknown length
 204228 205866: contig of 1639 bp in length
 205867 205966: gap of unknown length
 205967 207177: contig of 1211 bp in length
 207178 207277: gap of unknown length
 207278 208398: contig of 1121 bp in length

FEATURES
 SOURCE
 * 208399 208498: gap of unknown length
 * 208499 209551: contig of 1053 bp in length
 * 209552 209651: gap of unknown length
 * 209652 210698: contig of 1047 bp in length
 * 210699 210798: gap of unknown length
 * 210799 211875: contig of 1077 bp in length.
 Location/Qualifiers
 1..211875
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="CH240-1K23"
 30292..30341
 /estimated_length=50
 37666..37715
 /estimated_length=50
 53827..53876
 /estimated_length=50
 56441..56490
 /estimated_length=50
 67818..68524
 /estimated_length=707
 70610..71044
 /estimated_length=435
 73477..73526
 /estimated_length=50
 92742..92791
 /estimated_length=50
 96779..96828

Query Match 15.6%; Score 62.2; DB 12; Length 211875;
 Best Local Similarity 61.3%; Pred. No. 1.9e-09;
 Matches 100; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 8 CTCCTTCCTATATTATTTAGAGTTCTTGAATTAAGCTTAATTAATCTTGCGCTGACACAT 67
 Db 71851 CTCCTATAGTATGTTATTAACGACCTTGCGAGCATTAATTAATCTTGCAATGAATAC 71910
 Qy 68 GAGACTGATGATGTTAGCTTCTGGAAGTATGACGAGGCAAGTGAAGCTTACT 127
 Db 71911 AGAAGTATGATGATGCTTCTGGAAGTATGACGAGGCAAGTGAAGCTTACT 71970
 Qy 128 TTTTGTACTGTTAT 170
 Db 71971 TTTGCACTGACGGCTTTTATTAACATTAAGGATATACAA 72013

RESULT 8
 LOCUS AC122925/c 256751 bp DNA linear ROD 08-NOV-2003
 DEFINITION Mus musculus BAC clone RP23-9019 from 1, complete sequence.
 ACCESSION AC122925
 VERSION AC122925.3 GI:24476166
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Muridae; Murinae; Mus.
 1 (bases 1 to 256751)
 AUTHORS Harkins, R., Doeber, A. and Biellack, L.
 TITLE The sequence of Mus musculus BAC clone RP23-9019
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 256751)
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (28-MAY-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 256751)

AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 256751)
REFERENCE McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (02-NOV-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
6 (bases 1 to 256751)
REFERENCE Wilson, R.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 2, 2002 this sequence version replaced gi:23334975.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@wustl.edu
----- Summary Statistics
Center project name: M_BA009019

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:
The RPC1-23 BAC Library has been constructed by Kazuroyo Osegawa
and Minako Tateno in the laboratory of Pieter de Jong
(<http://www.chori.org>) from female C57BL/6J mouse kidney and/or
brain genomic DNA. The clone and detailed information can be
obtained from Research Genetics, Inc. (<http://www.resgen.com>) or
Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.
Location/Qualifiers
1..256751
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"
/clone="RPC23-9019"
/clone_lib="RPC1-23"
3239..3383
/rpt_family="A1u"
repeat_region
3957..4273
/rpt_family="MALR"
repeat_region
5279..5419
/rpt_family="ERVK"
repeat_region
5429..6070
/rpt_family="L1"
repeat_region
6071..6562
/rpt_family="ERVK"
repeat_region
8382..8633

repeat_region
8629..8747
/rpt_family="ERVK"
repeat_region
8748..9301
/rpt_family="ERVK"
repeat_region
9302..9467
/rpt_family="ERVK"
repeat_region
9523..10022
/rpt_family="ERVK"
repeat_region
10023..10101
/rpt_family="L1"
repeat_region
13484..13691
/rpt_family="B4"
repeat_region
15335..15475
/rpt_family="B4"
repeat_region
15614..15232
/rpt_family="L1"
repeat_region
17010..17083
/rpt_family="A1u"
repeat_region
18336..18496
/rpt_family="L1"
repeat_region
19428..19496
/rpt_family="TD"
repeat_region
19515..19733
/rpt_family="B4"
repeat_region
20439..20626
/rpt_family="L1"
21165..21733
/note="Sequence derived from PCR product of genomic DNA."
21885..22150
/note="Sequence derived from one plasmid subclone."
26874..27413
/rpt_family="ERVK"
27428..27573
/rpt_family="B4"
repeat_region
27933..28326
/rpt_family="MALR"
repeat_region
28769..29572
/rpt_family="ERVK"
repeat_region
29712..29928
/rpt_family="L1"
repeat_region
29928..30367
/rpt_family="L1"
repeat_region
30828..30886
/rpt_family="B4"
repeat_region
33156..33251
/rpt_family="ERVJ"
repeat_region
33401..33512
/rpt_family="L1"
repeat_region
34239..34610
/rpt_family="MALR"
repeat_region
37270..37700
/rpt_family="L1"
repeat_region
38033..38160
/rpt_family="B4"
repeat_region
38129..38179
/rpt_family="B4"
repeat_region
38681..38886
/rpt_family="L1"
repeat_region
39047..39479
/rpt_family="RMR6A"
repeat_region
39718..40075
/rpt_family="L1"
repeat_region
40191..40305
/rpt_family="L2"
repeat_region
42041..42191
/rpt_family="A1u"
repeat_region
42054..42228
/rpt_family="B4"
repeat_region
42960..43145
/rpt_family="L1"
repeat_region
47148..47311
/rpt_family="MIR"

REFERENCE
AUTHORS
1 (bases 1 to 152637)
Mizny,D,Marie, Metzker,M, Lee, Abramson,S, Adams,C, Alder,J, Allen,C, Allen,H, Albrooks,S, Amin,A, Anguiano,D, Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F, Biewala,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M, Bryant,N, Buhay,C, Burch,P, Burrill,K, Calderon,B, Cardenas,V, Carter,K, Cavazos,I, Caesar,H, Center,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D, Delgado,O, Denison,S, Deramo,C, Ding,Y, Dinh,H, Divya,K, Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K, Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falle,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P, Fraser,C,M, Gabisl,A, Ganta,R, Garcia,A, Garner,T, Garza,M, Gburegeorgie,B, Geier,K, Gill,R, Grady,M, Guerra,M, Guevara,W, Gunaratne,P, Haaland,W, Hamli,C, Hamilton,C, Hamilton,K, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hognes,M, Hollins,B, Howell,S, Hulyk,S, Hume,J, Idlebird,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Karpachy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C, Kowitz,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J, Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lourenhewa,L, Louisedge,H, Lorzado,R,J, Lu,X, Ma,J, Maheshwari,M, Mahindartne,M, Mahmoud,M, Malloy,K, Mangum,A, Mangun,B, Mapua,P, Martin,K, Martin,R, Martinez,B, Mathewley,S, McLeod,M,P, McNeill,T,Z, Meenen,E, Mlivosavljevic,A, Miner,G, Minja,G, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Mundasa,M, Murphy,M, Nair,L, Nankerville,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Nwankweme,O, Okonou,G, Olampunsgoon,A, Pal,S, Parks,K, Pasternak,S, Paul,H, Perez,A, Perez,L, Pfannkuch,C, Plopper,F, Polindexter,A, Popovic,D, Primus,E, Pu,L,L, Plazo,M, Quiroz,J, Rachlin,B, Reeves,K, Regier,M,A, Reigh,R, Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rives,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J, Sanders,W, Savery,G, Scherer,S, Scott,G, Shatman,S, Shen,H, Shetty,J, Shvartsbeyn,A, Sisson,I, Sitter,C,D, Smajs,D, Sneed,A, Sodergren,B, Song,X-Z, Sorelle,R, Sosa,J, Steimle,M, Strong,R, Sutton,A, Swatek,A, Taber,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Usamali,K, Valas,R, Vera,V, Villaseana,D, Walidron,L, Walker,B, Wang,J, Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F, Williams,G, Willson,R, Wleczek,R, Wooden,H, Worley,K, Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von Weinstock,G, and Gibbs,R.A.

TITLE
DIRECT Submission
Unpublished
2 (bases 1 to 152637)
Worley,K,C.
Direct Submission
Submitted (22-AUG-2004) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 152637)
Cow Genome Sequencing Consortium.
Direct Submission
Submitted (01-JUN-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jun 28, 2005 this sequence version replaced gi:51491771.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence configs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence configs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only configs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information

Center project name: FBOO
Center clone name: CH240-243C5
----- Summary Statistics

Assembly program: Atlas 3.0:
Consensus quality: 151855 bases at least Q40
Consensus quality: 152137 bases at least Q30
Consensus quality: 152302 bases at least Q20
Estimated insert size: 157598; sum-of-configs estimation
Estimated insert size: 142559; agarose-fp estimation
Quality coverage: 7x in Q20 bases; agarose-fp estimation
Quality coverage: 7x in Q20 bases; sum-of-configs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbankdraft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 configs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the configs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1019: contig of 1019 bp in length
* 1020 1119: gap of unknown length
* 1120 9204: contig of 8085 bp in length
* 9205 9304: gap of unknown length
* 9305 152637: contig of 143333 bp in length.

FEATURES

SOURCE

1. 152637
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-243C5"
1020. 1119
/estimated_length=unknown
9205. 9304
/estimated_length=unknown

ORIGIN

Query Match 14.9%; Score 59.6; DB 12; Length 152637;
Best Local Similarity 59.8%; Pred. No. 1.4e-08;
Matches 98; Conservative 1; Mismatches 65; Indels 0; Gaps 0;

QY 123 TACCTTTTGTACTGTTGATATATATATATACATACACACACACACACACA 182
DB 138340 TCCCTTTGAACAGGTGATCTTATATATATATATACACACACACACA 138281
QY 183 YATATTTTAACGCGACATATAATGTACACACACCTATTTCAAGTAAACTATCTCA 242
DB 138280 CATATATATATAAAAAATCATATATATATACATATCTTATATAGTTTATATAGACCTTTTAA 138221
QY 243 GAATTAAGACATATCTGCTATTTGATATTAACATAGAGTTAA 286
DB 138220 ATGTTAATAATATATATATATCTTTTATTTTGTGACATTTAA 138177

RESULT 11
AC113480 181457 bp DNA linear ROD 09-DEC-2005
LOCUS AC113480 Mus musculus chromosome 13, clone R23-309D8, complete sequence.
DEFINITION AC113480
AC113480.8 GI:83415361
VERSION

KEYWORDS
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Mus musculus chromosome 13, clone RP23-309D8
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 181457)

REFERENCE
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Chopel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., Mcwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (01-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS 3 (bases 1 to 181457)

REFERENCE
AUTHORS Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Chopel, Y., Collymore, A., Cook, A., Cooke, P., Cornu, B., Dearrellano, K., Diaz, J. S., Dodge, S., Doolley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Rachupka, A., Ramaamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (26-JUL-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS 4 (bases 1 to 181457)

REFERENCE
AUTHORS McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Rachupka, A., Ramaamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (20-SEP-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS 5 (bases 1 to 181457)

REFERENCE
AUTHORS Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Chopel, Y., Collymore, A., Cook, A., Cooke, P., Cornu, B., Dearrellano, K., Diaz, J. S., Dodge, S., Doolley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Rachupka, A., Ramaamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (09-DEC-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Dec 9, 2005 this sequence version replaced gi:75905661. All repeats were identified using RepeatMasker: <http://ftp.genome.washington.edu/RM/RepeatMasker.html>

FEATURES
source

 Center: Broad Institute of MIT and Harvard
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence.submissions@broad.mit.edu

 Project Information
 Center project name: L23717
 Center clone name: 309_D_8

 Location/Qualifiers
 1..181457
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="13"
 /map="13"
 /clone="RP23-309D8"
 /clone_lib="RPCI-23 Female Mouse BAC"
 complement(1..91)
 /rpt_family="B1_MM"
 repeat_region
 314..2008
 /rpt_family="ORR1A-int"
 repeat_region
 2007..2396
 /rpt_family="ORR1B2"
 repeat_region
 2437..2566
 /rpt_family="B1P"
 repeat_region
 2572..2625
 /rpt_family="GA-rich"
 repeat_region
 3350..3635
 /rpt_family="Lx7"
 repeat_region
 3636..3829

1	694	: conf: of 694 bp	in	length
695	794	: gap of 100 bp	in	length
795	1505	: conf: of 711 bp	in	length
1506	1605	: gap of 100 bp	in	length
1606	2278	: conf: of 673 bp	in	length
2279	2378	: gap of 100 bp	in	length
2379	3043	: conf: of 665 bp	in	length
3044	3143	: gap of 100 bp	in	length
3144	3831	: conf: of 688 bp	in	length
3832	3931	: gap of 100 bp	in	length
3932	4624	: conf: of 693 bp	in	length
4625	4724	: gap of 100 bp	in	length
4725	5400	: conf: of 696 bp	in	length
5421	5520	: gap of 100 bp	in	length
5521	6248	: conf: of 728 bp	in	length
6249	6349	: conf: of 711 bp	in	length
6349	7059	: conf: of 711 bp	in	length
7060	7159	: gap of 100 bp	in	length
7160	7841	: conf: of 682 bp	in	length
7842	7941	: gap of 100 bp	in	length
7942	8643	: conf: of 702 bp	in	length
8644	8743	: gap of 100 bp	in	length
8744	9460	: conf: of 717 bp	in	length
9461	9560	: gap of 100 bp	in	length
9561	10260	: conf: of 700 bp	in	length
10261	10360	: gap of 100 bp	in	length
10361	11040	: conf: of 680 bp	in	length
11041	11140	: gap of 100 bp	in	length
11141	11817	: conf: of 677 bp	in	length
11818	11917	: gap of 100 bp	in	length
11918	12602	: conf: of 685 bp	in	length
12603	12702	: gap of 100 bp	in	length
12703	13407	: conf: of 705 bp	in	length
13408	13507	: gap of 100 bp	in	length
13508	14225	: conf: of 718 bp	in	length
14226	14325	: gap of 100 bp	in	length
14326	15026	: conf: of 701 bp	in	length
15027	15126	: gap of 100 bp	in	length
15127	15827	: conf: of 701 bp	in	length
15828	15927	: gap of 100 bp	in	length
15928	16602	: conf: of 675 bp	in	length
16603	16702	: gap of 100 bp	in	length
16703	17411	: conf: of 709 bp	in	length
17412	17511	: gap of 100 bp	in	length
17512	18207	: conf: of 696 bp	in	length
18208	18307	: gap of 100 bp	in	length
18308	19012	: conf: of 705 bp	in	length
19013	19112	: gap of 100 bp	in	length
19113	19775	: conf: of 663 bp	in	length
19776	19875	: gap of 100 bp	in	length
19876	20560	: conf: of 685 bp	in	length
20561	20660	: gap of 100 bp	in	length
20661	21325	: conf: of 665 bp	in	length
21326	21425	: gap of 100 bp	in	length
21426	22119	: conf: of 694 bp	in	length
22120	22219	: gap of 100 bp	in	length
22220	22913	: conf: of 694 bp	in	length
22914	23013	: gap of 100 bp	in	length
23014	23725	: conf: of 712 bp	in	length
23726	23825	: gap of 100 bp	in	length
23826	24529	: conf: of 704 bp	in	length
24530	24629	: gap of 100 bp	in	length
24630	25308	: conf: of 679 bp	in	length
25309	25408	: gap of 100 bp	in	length
25409	26095	: conf: of 687 bp	in	length
26096	26195	: gap of 100 bp	in	length

Query Match	14.8%	Score 59;	DB 12;	Length 64239;
Best Local Similarity	56.6%	Pred. No. 1.9e-08;		
Matches 107;	Conservative 1;	Mismatches 81;	Indels 0;	Gaps 0

Oy	128	TTTGTGACGTGTTGGATATATATATTAATACATACACACACACACACACATAT	187
Db	23077	TTTTATCTTTTTTATTAACATATATATATATATATACACACACACACACACACAT	23138
Oy	188	TTTAACTGGGACATTAATAGTACACGACCTATTCAAAGTAAAACTACTATCTCAGANGT	247
Db	23137	ACATACATGTGGAAAGCCCAATCTATCCCTGTTCCCAATCCAAACATATTTTTTTTTTT	23196
Oy	248	AAAGACATTAATCCTGCTATTTGTCATTAACTAGACAGTTAAAGTGCCTAACCTTTATGCACT	307
Db	23197	TGAGACAGAGATCTTGCTCTGTGSCCAGGCTGGCGGTCAGTGGCAAAATCTCCGCTCACT	23256
Oy	308	TCAGTTTCC	316
Db	23257	GCAACTTCC	23265

[illegible]

```

REFERENCE
AUTHORS
TITLE
JOURNAL

5 (bases 1 to 104485)
Worley,K.C.
Direct Submission
Submitted (01-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
Baylor Plaza, Houston, TX 77030, USA
On Jun 25, 2002 this sequence version replaced gi:21327308.
COMMENT
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:
STSs are identified using EPCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
http://www.hgsc.bcm.tmc.edu:8086/quality.info/genbank.annotation.htm.

FEATURES
source
Location/Qualifiers
1..104485
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-12K11"
/complement(5..2005)
/note="overlaps bases 1..2001 of clone AC128686"
/function="clone overlap"
repeat_region
complement(78..175)
/rpt_family="MER41B"
complement(173..574)
/rpt_family="MER41A"
593..635
/rpt_family="AT_rich"
698..838
/rpt_family="HAL1"
839..1149
/rpt_family="ALUy"
1150..1271
/rpt_family="HAL1"
1277..1490
/rpt_family="L1M4"
complement(1750..1792)
/rpt_family="L1MA10"
1793..2096
repeat_region

```


LOCUS AC027755 147328 bp DNA linear HTG 02-MAY-2000
 DEFINITION Homo sapiens chromosome 3 clone RP11-594C8 map 3, WORKING DRAFT
 ACCESSION AC027755
 VERSION AC027755.2 GI:7677887
 KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 REFERENCE 1 (bases 1 to 147328)
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 TITLE Homo sapiens chromosome 3, clone RP11-594C8
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 147328)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
 Boguslavsky, L., Bouckgalter, B., Brown, A., Burkett, G.,
 Campopiano, A., Castle, A., Choquel, Y., Colangelo, M., Collins, S.,
 Collamore, A., Cooke, P., DeArrelano, K., Dewar, K., Diaz, J. S.,
 Dodge, S., Domingo, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kam, L., Karacas, A.,
 Klein, J., Larocque, K., Lamazares, R., Landers, T., Lenoczky, J.,
 Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,
 McCarthy, M., McEwan, P., McGuire, A., McKernan, K., McPheters, R.,
 Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenaga, V., Morrow, J.,
 Murphy, T., Naylor, T. M., Norman, C. H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N., P.,
 Pisaní, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Teefaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.
 DIRECT SUBMISSION
 Submitted (01-APR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 2, 2000 this sequence version replaced gi:7382581.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 TITLE
 JOURNAL
 COMMENT
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submission@genome.wi.mit.edu
 Project Information
 Center project name: 594_C_8
 Center clone name: L8252
 Summary Statistics
 Sequencing vector: M13; M77815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 140084 bases at least Q40
 Consensus quality: 143890 bases at least Q30
 Consensus quality: 145141 bases at least Q20
 Insert size: 150000; agarose-fp
 Insert size: 145928; sum-of-contigs
 Quality coverage: 4.6 in Q20 bases; agarose-fp
 Quality coverage: 4.7 in Q20 bases; sum-of-contigs
 NOTES: This is a 'working draft' sequence. It currently
 consists of 15 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 674: contig of 674 bp in length

FEATURES	SOURCE
675	774: gap of 100 bp
775	3066: contig of 2292 bp in length
3067	3166: gap of 100 bp
3167	6613: contig of 3447 bp in length
6614	6713: gap of 100 bp
6714	11352: contig of 4639 bp in length
11353	11452: gap of 100 bp
11453	17821: contig of 6369 bp in length
17822	17921: gap of 100 bp
17922	25702: contig of 7781 bp in length
25703	25802: gap of 100 bp
25803	32656: contig of 6854 bp in length
32657	32756: gap of 100 bp
32757	36892: contig of 7136 bp in length
36893	39992: gap of 100 bp
39993	48004: contig of 8012 bp in length
48005	48104: gap of 100 bp
48105	56776: contig of 8572 bp in length
56777	56776: gap of 100 bp
56778	68565: contig of 11789 bp in length
68566	68665: gap of 100 bp
68666	81186: contig of 12521 bp in length
81187	81286: gap of 100 bp
81287	97090: contig of 15804 bp in length
97091	97190: gap of 100 bp
97191	119119: contig of 21929 bp in length
119120	119219: gap of 100 bp
119220	147328: contig of 28109 bp in length.
147328	Location/Qualifiers
1..147328	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="3"
	/map="3"
	/clone="RP11-594C8"
	/clone_lib="RP11 Human Male BAC"
1..674	/note="assembly_fragment"
	clone_end:77
	vector_side:rightc"
675..774	/estimated_length=100
775..3066	/note="assembly_fragment"
3067..3166	/estimated_length=100
3167..6613	/note="assembly_fragment"
6614..6713	/estimated_length=100
6714..11352	/note="assembly_fragment"
11353..11452	/estimated_length=100
11453..17821	/note="assembly_fragment"
17822..17921	/estimated_length=100
17922..25702	/note="assembly_fragment"
25703..25802	/estimated_length=100
25803..32656	/note="assembly_fragment"
32657..32756	/estimated_length=100
32757..36892	/note="assembly_fragment"
36893..39992	/estimated_length=100
39993..48004	/note="assembly_fragment"


```

misc_feature      1..1056
                  /note="assembly_fragment"
gap               1057..1156
                  /estimated_length=100
misc_feature      1157..14629
                  /note="assembly_fragment"
gap               14630..14729
                  /estimated_length=100
misc_feature      14730..15865
                  /note="assembly_fragment"
gap               15866..15965
                  /estimated_length=100
misc_feature      15966..19176
                  /note="assembly_fragment"
gap               19177..19276
                  /estimated_length=100
misc_feature      19277..12888
                  /note="assembly_fragment"
gap               12889..12988
                  /estimated_length=100
misc_feature      12989..15953
                  /note="assembly_fragment"
gap               15954..16053
                  /estimated_length=100
misc_feature      16054..121676
                  /note="assembly_fragment"
gap               121677..121776
                  /estimated_length=100
misc_feature      121777..127872
                  /note="assembly_fragment"
gap               127873..127972
                  /estimated_length=100
misc_feature      127973..137824
                  /note="assembly_fragment"
gap               137825..137924
                  /estimated_length=100
misc_feature      137925..147260
                  /note="assembly_fragment"
gap               147261..147360
                  /estimated_length=100
misc_feature      147361..160391
                  /note="assembly_fragment"
gap               160392..160491
                  /estimated_length=100
misc_feature      160492..184436
                  /note="assembly_fragment"
gap               184437..184536
                  /estimated_length=100
misc_feature      184537..114093
                  /note="assembly_fragment"
gap               114094..114193
                  /estimated_length=100
misc_feature      114194..114851
                  /note="assembly_fragment"
ORIGIN
Query Match      14.7% Score 58.6; DB 12; Length 148851;
Best Local Similarity 57.5%; Pred. No. 3e-08;
Matches 103; Conservative 1; Mismatches 75; Indels 0; Gaps 0;

```

```

Db               27024 TATACATATATATGATTTATATACATTTAGCGCTGAAGTAGAAGAACTGTAGACATCAGC 26966
RESULT 17
AC108728/c
LOCUS
DEFINITION
AC108728      303189 bp    DNA      linear      HTG 09-MAY-2002
Homo sapiens chromosome 3 clone RP11-594C8, WORKING DRAFT SEQUENCE,
18 unordered pieces.
AC108728
AC108728.2 GI:20335976
VERSION
HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
1 (bases 1 to 303189)
Muzny D.M., Adams C., Adio-Oduola B., Alf-Osman F.R., Allen C.,
Albrooks S.L., Amaralunge H.C., Are J.R., Ayele M., Banks T.,
Barbarta J., Benton J., Bimane K., Blankenburg K., Bonnin D.,
Bouck J., Bowle S., Brileva M., Brown E., Brown M., Bryant N.P.,
Bunay C., Burch P., Burkett C., Burrell K.L., Byrd N.C.,
Carron T.F., Carter M., Cavazos S.R., Chacko J., Chavez D.,
Chen G., Chen R., Chen Z., Chowdhury I., Christopoulos C.,
Cleveland C.D., Cox C., Coyle M.D., Dathorne S.R., David R.,
Dayila M.L., Davis C., Davy-Carroll J., Dedrich D.A.,
DeJaney K.R., Delgado O., Denn A.L., Ding Y., Dinb H.H.,
Douthwaite K.J., Draper H., Dugan-Rocha S., Durbin K.J.,
Earhart C., Edgar D., Edwards C.C., Elhaj C., Escoto M.,
Falls T., Ferraguto D., Flagg N., Ford J., Foster P., Frantz P.,
Gabisi A., Gao J., Garcia A., Garner T., Garza N., Gill R.,
Gorrell J., H., Guevara W., Gunaratne P., Hale S., Hamilton K.,
Harris C., Harris K., Hart M., Haylak P., Hawes A., Hernandez J.,
Hernandez O., Hodgson A., Hognes M., Holloway C., Hollins B.,
Homai F., Howard S., Huber J., Hulky S., Hume J., Jackson L.E.,
Jacobson B., Jia Y., Johnson R., Jolivet S., Joudah S.,
Karlsson E., Kelly S., Khan U., King L., Korvah J., Kovar C.,
Kratovic J., Kureshi A., Landry N., Leal B., Lewis L.C., Lewis L.,
Li J., Li Z., Lichtarge O., Lieu C., Liu J., Liu W., Louised H.,
Lozado R.J., Lu X., Lucier A., Lucier R., Luna R., Ma J.,
Maheshwari M., Mapa P., Martin R., Martindale A., Martinez E.,
Massey G., Mashiney E., McLeod M.P., Meador M., Mei G., Metzger M.,
Miner G., Miner Z., Mitchell T., Mohabbat K., Morgan M., Morris S.,
Moser M., Neal D., Newton J., Newton N., Nguyen A., Nguyen N.,
Nguyen N., Nickerson E., Nwokenko S., Ogil M., Okunodu G.,
Oregon N., Oviedo R., Pace A., Payton B., Peery J., Perez L.,
Peters L., Pickens R., Primus E., Pu L.L., Quiles M., Ren Y.,
Rivers M., Rojas A., Rojokan I., Rolfe M., Ruiz S., Savery G.,
Scherer S., Scott G., Shen H., Shoostari N., Sisson I.,
Sodergren E., Sonaite T., Sparks A., Stanley H., Stone H.,
Sutton A., Svatek A., Tabor P., Tamezisa A., Tamezisa K., Tang H.,
Tansley J., Taylor C., Taylor T., Telford B., Thomas N., Thomas S.,
Umani K., Vasquez L., Vera V., Villalob D., Vinson R., Wang O.,
Wang S., Ward-Moore S., Warren R., Washington C., Watlington S.,
Williams G., Williamson A., Wleciyk R., Wooden S., Worley K.,
Wu C., Wu Y., Wu Y.F., Zhou J., Zortilla S., Nelson D.,
Weinstock G. and Gbbs R.
Direct Submission
Unpublished
2 (bases 1 to 303189)
Worley K.C.
Direct Submission
Submitted (31-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 303189)
Worley K.C.
Direct Submission
Submitted (09-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
COMMENT
On Apr 28, 2002 this sequence version replaced gi:18449858.

```

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HDVG
Center clone name: RP11-594C8
----- Summary Statistics
Sequencing vector: Plasmid:
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 295424 bases at least Q40
Consensus quality: 300495 bases at least Q30
Consensus quality: 303167 bases at least Q20
Estimated insert size: 152199, sum-of-coverage estimation
Quality coverage: 6x in Q20 bases; sum-of-coverage estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 3465: contig of 3465 bp in length
* 3466 3566: gap of unknown length
* 3566 7534: contig of 3969 bp in length
* 7534 7535: gap of unknown length
* 7535 13994: contig of 6360 bp in length
* 13994 14094: gap of unknown length
* 14094 22000: contig of 7906 bp in length
* 22001 22101: gap of unknown length
* 22101 30722: contig of 8622 bp in length
* 30723 30823: gap of unknown length
* 30823 42634: contig of 11812 bp in length
* 42635 42734: gap of unknown length
* 42734 54975: contig of 12241 bp in length
* 54975 55075: gap of unknown length
* 55075 67039: contig of 11964 bp in length
* 67040 67139: gap of unknown length
* 67139 79747: contig of 12608 bp in length
* 79748 79748: gap of unknown length
* 79748 91971: contig of 11214 bp in length
* 91972 91972: gap of unknown length
* 91972 108785: contig of 16714 bp in length
* 108786 108885: gap of unknown length
* 108886 130010: contig of 21125 bp in length
* 130011 130110: gap of unknown length
* 130111 146359: contig of 16249 bp in length
* 146360 146459: gap of unknown length
* 146460 171632: contig of 25173 bp in length
* 171633 171732: gap of unknown length
* 171733 194849: contig of 23117 bp in length
* 194850 194949: gap of unknown length
* 194950 223128: contig of 28179 bp in length
* 223129 223228: gap of unknown length
* 223229 259005: contig of 35777 bp in length
* 259006 259105: gap of unknown length
* 259106 303189: contig of 44084 bp in length.

FEATURES
SOURCE
1..303189
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-594C8"
3466..3565
/estimated_length=unknown
7535..7634

```

gap	/estimated_length=unknown
13995. .14094	
gap	/estimated_length=unknown
22001. .22100	
gap	/estimated_length=unknown
30723. .30822	
gap	/estimated_length=unknown
42635. .42734	
gap	/estimated_length=unknown
54976. .55075	
gap	/estimated_length=unknown
67040. .67139	
gap	/estimated_length=unknown
79748. .79847	
gap	/estimated_length=unknown
91972. .92071	
gap	/estimated_length=unknown
108786. .108885	
gap	/estimated_length=unknown
130011. .130110	
gap	/estimated_length=unknown
146360. .146459	
gap	/estimated_length=unknown
171633. .171732	
gap	/estimated_length=unknown
194850. .194949	
gap	/estimated_length=unknown
223129. .223228	
gap	/estimated_length=unknown
259006. .259105	
gap	/estimated_length=unknown

	Breast Match	Similarity	14.7%	Score 58.6;	DB 12;	Length 303189;	
	Query Local	Similarity	57.5%	Pred. No. 3.3e-08;			
	Matches 103;	Conservative	1;	Mismatches 75;	Indels 0;	Gaps 0;	
Oy	110	GGCAGTGTGAGGCCTACCTTTTGTCATGTTTGATATATATATATATATATATACATACACACACAC 169					
Dd	204738	GGCGAGTGTCAAAATATCCAAACTATATCATTCATTCTATATATATATATACACACACACACAC 204679					
Oy	170	AACACACACACACACATATATTTTAACCTCGGACACTAATATGTCACACACCCTTTCAAAGTAAA 229					
Dd	204678	ACACACTCACACACACATGAATGATATATACATACATACATATATATATGAACGATATATATATATAT 204619					
Oy	230	AACCTACTATCTCCAGAAGTAAAGACATATCCGTCATTTGTCTATTAACCTAGCATTTAACG 288					
Dd	204618	TATACATATATATATGAATTTATATACTATTAGCTGAAAAAGTGAACAACCTGTAGAGCATCAGC 204560					
RESULT 18							
LOCUS	AC021491						
DEFINITION	Homo sapiens chromosome 17, clone RP11-94C24, complete sequence.						
ACCESSION	AC021491						
VERSION	AC021491.16 GI:19747233						
KEYWORDS	HTG.						
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.						
REFERENCE	1 (bases 1 to 175262) Britten,B., Linton,L., Nusbaum,C. and Landier,E. Homo sapiens chromosome 17, clone RPl1-94C24 Unpublished						
AUTHORS	2 (bases 1 to 175262) Britten,B., Linton,L., Nusbaum,C., Landier,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Beckely,R., Bede,F., Boquerville,K.L., Boukhalter,B., Brown,A., Burkett,G., Castle,A., Chapel,Y., Colangelo,M., Collins,S., Collimore,A., Cooke,P., Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenebor,J., Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,						
JOURNAL							
REFERENCE							
AUTHORS							

Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kam, L., Karatas, A., Klein, J., Landers, T., Lehoczy, J., Levine, R., Lien, C., Liu, G., Locke, K., Macdonald, P., Margulis, N., McEwan, P., McGuirk, A., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirelli, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J., Zimmer, A., and Zody, M.

TITLE

Direct Submission

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome

REFERENCE

Research, 320 Charles Street, Cambridge, MA 02141, USA

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fato, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Milnova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Ntubu, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Direct Submission

Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome

JOURNAL

Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

4 (bases 1 to 175262)

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fato, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Milnova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Ntubu, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Direct Submission

Submitted (27-MAR-2002) Whitehead Institute/MIT Center for Genome

JOURNAL

Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Mar 27, 2002 this sequence version replaced gi:19703305.

All repeats were identified using RepeatMasker:
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

FEATURES

source

Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L5368
 Center clone name: 94_C_24

Only the first 175.3 kilobases of this clone are being submitted.
 The remainder overlaps accession number AC004707 [WICGR project L36].

repeat_region	Location/Qualifiers
repeat_region	1..175262
repeat_region	/organism="Homo sapiens"
repeat_region	/mol_type="genomic DNA"
repeat_region	/db_xref="taxon:9606"
repeat_region	/chromosome="17"
repeat_region	/map="17"
repeat_region	/clone_id="RP11-94C24"
repeat_region	/clone_lib="RPC1-11 Human Male BAC"
repeat_region	511..612
repeat_region	/rpt_family="CAGAGA)n"
repeat_region	921..950
repeat_region	/rpt_family="TTTTTG)n"
repeat_region	966..989
repeat_region	/rpt_family="(TGA)n"
repeat_region	complement(11770..2004)
repeat_region	/rpt_family="MER102b"
repeat_region	complement(2005..2293)
repeat_region	/rpt_family="AluYo"
repeat_region	complement(2294..2392)
repeat_region	/rpt_family="MER102b"
repeat_region	complement(3017..3217)
repeat_region	/rpt_family="MIR"
repeat_region	3204..3280
repeat_region	/rpt_family="MIR"
repeat_region	3915..4060
repeat_region	/rpt_family="MER5A"
repeat_region	complement(4068..4249)
repeat_region	/rpt_family="MIR"
repeat_region	complement(7603..7696)
repeat_region	/rpt_family="L2"
repeat_region	8974..9095
repeat_region	/rpt_family="C-rich"
repeat_region	10187..10360
repeat_region	/rpt_family="MIR"
repeat_region	10401..10487
repeat_region	/rpt_family="L1MA7"
repeat_region	10488..10767
repeat_region	/rpt_family="AluX"
repeat_region	10768..11186
repeat_region	/rpt_family="L1MA7"
repeat_region	11187..11461
repeat_region	/rpt_family="AluX"
repeat_region	11465..11491
repeat_region	/rpt_family="(GAAA)n"
repeat_region	11492..12020
repeat_region	/rpt_family="L1MA7"
repeat_region	12021..12047
repeat_region	/rpt_family="AT-rich"
repeat_region	12050..12270
repeat_region	/rpt_family="MIR"
repeat_region	complement(12304..12404)
repeat_region	/rpt_family="MIR"
repeat_region	complement(14791..14995)
repeat_region	/rpt_family="MIR"
repeat_region	19566..19602
repeat_region	/rpt_family="(CA)n"
repeat_region	19732..19807
repeat_region	/rpt_family="MIR"
repeat_region	complement(22107..22298)
repeat_region	/rpt_family="MIR"
repeat_region	22772..22901
repeat_region	/rpt_family="MIR"

Query Match	14.3%	Score 57.2;	DB 5;	Length 175262;
Best Local Similarity	55.0%	Pred. No. 9.2e-08;		
Matches 110; Conservative	1;	Mismatches 89;	Indels 0;	Gaps 0;

TITLE
Direct SubMISSION
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 15, 2001 this sequence version replaced g1:11225458.
COMMENT
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

```

Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information -----
Center project name: l7968
Center clone name: 262_F21
----- Summary Statistics -----
Sequencing vector: M13; M7815; 4% of reads
Sequencing vector: Plasmid; n/a; 96% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 146016 bases at least Q40
Consensus quality: 147802 bases at least Q30
Consensus quality: 148377 bases at least Q20
Insert size: 150000; agarose-fp
Insert size: 148579; sum-of-contigs
Quality coverage: 6.5 in Q20 bases; agarose-fp
Quality coverage: 6.5 in Q20 ba.

```


REFERENCE 2 (bases 1 to 112727)
AUTHORS Do,T., Drabek,B. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (30-OCT-2001) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE 3 (bases 1 to 112727)
AUTHORS Do,T., Drabek,B. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE 4 (bases 1 to 112727)
AUTHORS Do,T., Drabek,B. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE 5 (bases 1 to 112727)
AUTHORS Do,T., Drabek,B. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE 6 (bases 1 to 112727)
AUTHORS Do,T., Drabek,B. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

COMMENT On Mar 9, 2004 this sequence version replaced gi:42734074.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKMOR

FEATURES
source
location/Qualifiers
1. 112727
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129/Sv"
/db_xref="taxon:10090"
/map="5"
/clone="cct-128f23"
/clone_lib="c1tbc7 mouse BAC library"

ORIGIN
Query Match 14.1%; Score 56.2; DB 6; Length 112727;
Best Local Similarity 70.9%; Pred. No. 1.9e-07;
Matches 73; Conservative 1; Mismatches 29; Indels 0; Gaps 0;

QY 90 TCTGGAAGTAGTATGAGGGCAAGGCTTCTTTGTAAGTCTTGATATATAT 149
Db 82837 TCTTTAAAAAATCCAAACAAAGCCAGGCTTACCAAGACCTTGCTCAAAATAT 82896

QY 150 ATATACATACACACACACACACACACACACACACACATATATTTCA 192
Db 82897 ATATACACACACACACACACACACACACACACATATATTTCA 82939

RESULT 23
AC060796/c 118646 bp DNA linear PRI 19-FEB-2002
LOCUS Homo sapiens chromosome 17, clone CTD-2245L18, complete sequence.
ACCESSION AC060796 GI:18702411
VERSION AC060796.8
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 118646)

AUTHORS Birren,B., Linton,L., Nuebaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone CTD-2245L18
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 118646)
AUTHORS Birren,B., Linton,L., Nuebaum,C., Lander,E., Abrahams,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Basleten,V., Bada,F.,
Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G.,
Campiano,A., Casale,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Glnde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hago,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kamm,L., Karakas,A.,
Klein,J., Laroque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Liu,C., Liu,G., Locke,K., Riley,R., Rogov,P., Rothman,D.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,
Meldrum,J., Menus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teafaye,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 118646)
AUTHORS Birren,B., Linton,L., Nuebaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Basleten,V., Boguslavsky,L., Boukhalter,B.,
Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hago,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karakas,A., Kells,J.C., Laroque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
McClennan,C., MacDonald,P., Major,J., Margulis,N., Matchew,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrum,J.,
Menus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Punthang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R.,
Seaman,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
Straus,N., Subramanian,A., Talamas,J., Teafaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (11-SEP-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 118646)
REFERENCE 1 (bases 1 to 118646)
AUTHORS Birren,B., Linton,L., Nuebaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Basleten,V., Boguslavsky,L., Boukhalter,B.,
Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hago,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karakas,A., Kells,J.C., Laroque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., McClennan,C.,
MacDonald,P., Major,J., Margulis,N., Matchew,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Menus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Punthang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R.,
Seaman,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
Teafaye,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Only the last 118.6 kilobases of this clone are being submitted. The remainder overlaps accession number AC006263 [WICR project L536].

```

RES
source
location/Qualifiers
1. 118646
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/cclone="CTD-2245L18"
/cclone_lib="CTD Human BAC"
204. 497
/rpt_family="AluSp"
complement(977. 1101)
/rpt_family="MIR3"
complement(1507. 1623)
/rpt_family="L2"
1736. 1779
/rpt_family="DCGRn"
complement(1198. 1860)
/rpt_family="L2"
complement(1957. 2036)
/rpt_family="MLT1A"
complement(2154. 2327)
/rpt_family="L2"
2920. 3196
/rpt_family="MLT1A"
complement(3472. 3554)
/rpt_family="MIR3"
3924. 3946
/rpt_family="AT_rich"
complement(5220. 5523)
/rpt_family="AluSx"
complement(5650. 5946)
/rpt_family="AluY"
7304. 7601
/rpt_family="AluSd"
complement(8436. 8743)
/rpt_family="MARNA"
8823. 8858
/rpt_family="UTRA"
complement(8862. 8966)
/rpt_family="FLAM_A"
9002. 9236
/rpt_family="AluSx"
9311. 9359
/rpt_family="FAM"
9435. 9491
/rpt_family="L2"
complement(10233. 10381)
/rpt_family="MIR"

```

```
repeat_region complement (12121. , 12151)
/rpt_family="MLT2B1"
12152. .12199
/rpt_family=" (TG)n"
repeat_region complement (12200. , 12714)
/rpt_family="MLT2B1"
12811. .12894
/rpt_family="L2"
repeat_region complement (12962. , 13163)
/rpt_family="MER107"
complement (13737. , 13999)
/rpt_family="AluDb"
complement (14674. , 14970)
/rpt_family="AluSx"
15059. .15233
/rpt_family="L2"
15234. .15447
/rpt_family="AluSg/x"
15448. .15481
/rpt_family=" (TG)n"
15482. .15551
/rpt_family="L2"
15643. .15721
/rpt_family="L2"
complement (15722. , 15982)
/rpt_family="AluSx"
15983. .16190
/rpt_family="L2"
16411. .16435
/rpt_family="AT_rich"
16666. .16773
/rpt_family="L1ME4A"
17002. .17257
/rpt_family="Tigger7"
complement (17448. , 17648)
/rpt_family="MLT1J"
17964. .18282
/rpt_family="AluSc"
complement (18531. , 18833)
/rpt_family="AluJo"
19341. .19635
/rpt_family="AluSx"
20003. .20285
```

Query Match	14.1%;	Score 56.2;	DB 5;	Length 118646;
Best Local Similarity	63.0%;	Pred. No. 1.9e-07;		
Matches 85;	Conservative 1;	Mismatches 49;	Indels 0;	Gaps 0

[illegible]

AC156029	LOCUS	DEFINITION	AC156029	235471 bp	DNA	linear	ROD 23-APR-2005
AC156029		Mus musculus BAC clone RP23-4208 from chromosome 5, complete sequence.	AC156029				
AC156029.2		GI:61807066					
KEYWORDS		HTG.					
SOURCE		Mus musculus (house mouse)					
ORGANISM		Mus musculus					
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;					

Puzzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Rella, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P.,
Rivers, C., Rodkey, T., Rojase, A., Rose, M., Rose, R., Ruiz, S.J.,
Sadders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, I., Shivatsbeyn, A., Sisson, I., Slitter, C.D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Taboz, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umant, K.,
Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weise, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.

Unpublished

Direct Submission

2 (bases 1 to 236644)

Worley, K.C.

Direct Submission

Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 236644)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:24818703.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center -----

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information -----

Center project name: GSZF

Center clone name: CH230-33N6

----- Summary Statistics -----

Assembly program: Atlas 3.0;
Consensus quality: 228522 bases at least Q40
Consensus quality: 229930 bases at least Q30
Consensus quality: 231245 bases at least Q20
Estimated insert size: 243865; sum-of-contigs estimation
Quality coverage: 12x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html),
NOTE: This is a 'working draft' sequence. It currently
consists of 6 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 229587: contig of 229587 bp in length

* 229588 229687: gap of unknown length

* 229688 231130: contig of 1443 bp in length

* 231131 231230: gap of unknown length

* 231231 232260: contig of 1030 bp in length

	*	232261	23360:	gap of unknown length
	*	232361	233664:	contig of 1304 bp in length
	*	233665	233764:	gap of unknown length
	*	233765	235149:	contig of 1385 bp in length
	*	235150	235249:	gap of unknown length
	*	235250	236644:	contig of 1395 bp in length.
FEATURES				
source			location/Qualifiers	
			1..236644	
			/organism="Rattus norvegicus"	
			/mol_type="genomic DNA"	
			/db_xref="taxon:10116"	
			/clone="CH230-33W6"	
			1..1787	
misc_feature			/note="wgs_end_extension clone_end:T7"	
			1838..3553	
misc_feature			/note="wgs_end_extension clone_end:T7"	
			4795..5607	
misc_feature			/note="clone_boundary clone_end:T7 site:ECORI"	
			end sequence:BH356361"	
			complement(225661..226474)	
			/note="clone_boundary clone_end:Sp6 site:ECORI"	
			end sequence:BH356372"	
			226880..229587	
misc_feature			/note="wgs_end_extension clone_end:Sp6"	
			229588..229687	
gap			/estimated_length=unknown	
			231131..231230	
gap			/estimated_length=unknown	
			232261..232360	
gap			/estimated_length=unknown	
			233665..233764	
gap			/estimated_length=unknown	
			235150..235249	
gap			/estimated_length=unknown	
ORIGIN				
Query Match			14.1%; Score 56.2; DB 12;	Length 236644;
Best Local Similarity			59.1%; Pred. No. 2.1e-07;	
Matches			94; Conservative	1; Mismatches 64; Indels 0; Gaps 0;
OY		142	ATATATTATTTACATACACACACACACACACACACACACATATTTTAACTCGGGACA	201
Db		78613	ACATATTTACATACACACACACACACACACACACATACGGAAGACTGAA	7855
OY		202	TAAATGTACACCACCTTCATAAAGTAATAAACCTCATCTCAGAAGTAAGACATTAATCT	261
Db		78553	AGAAATGGACCTAATATTTAAACTAAATAATTTTTTAAAAATTTAAAATTAATTAAT	7849
OY		262	GCTATTGTCAATTAAGTAGAGCTTAAGCTGCCTAACCTTT	300
Db		78493	TCAAAGGAACATTTTACTTCAGGAACCTGTGAATTTT	78455
RESULT 26				
ACI33731/c			239980 bp	DNA linear HTG 15-NOV-2002.
LOCUS				
DEFINITION			Rattus norvegicus clone CH230-1116,	*** SEQUENCING IN PROGRESS ***
ACCESSION			6 unordered pieces.	
VERSION			ACI33731	
KEYWORDS			HTG; HTGS PHASER; HTGS DRAFI; HTGS_ENRICHED.	
SOURCE			Rattus norvegicus (Norway rat)	
ORGANISM			Rattus norvegicus	
			Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Rattus.	

REFERENCE

1 (bases 1 to 239980)

Wuzy,D, Marie, Metzker,M, Lee, Abramzon,S, Adams,C, Alder,J, Allen,C, Allen,H, Aishbrook,S, Amin,A, Anguiano,D, Ayala-Bechechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F, Bialwal,D, Blair,J, Blankenburg,K, Blyth,P, Brown,M, Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E, Cardenas,V, Carter,K, Cavazos,I, Ceasar,H, Centen,A, Chaco,J, Chavez,D, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L, Davila,M, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D, Delgado,O, Denison,S, Deramo,C, Ding,Y, Dinh,H, Divya,K, Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K, Egan,A, Escotto,M, Eugene,C, Evans,C, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P, Fraser,C, Gabisi,A, Gant,R, Garcia,A, Garner,T, Garza,M, Gatregeorgis,E, Geer,K, Gill,R, Grady,M, Guerra,M, Guevara,W, Gunaratne,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hladun,S, Hodgson,A, Hoques,M, Hollins,B, Howells,S, Huliy,S, Hume,J, Idlebird,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Karpachy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovat,C, Kowls,C, Kraft,C, Lebow,H, Levan,D, Lewis,L, Li,Z, Liu,J, Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lorenshewa,L, Loulsegad,H, Lozada,R, Lu,X, Ma,J, Maheshwari,M, Mahindartine,M, Mahmoud,M, Malloy,K, Mangum,A, Mangum,B, Mapus,P, Martin,K, Martin,R, Martinez,E, Mawhiley,S, McLeod,M, McNeill,T, Z, Meenen,E, Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Mundasa,M, Murphy,M, Nair,L, Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Nwokilemech,O, Okwunonu,G, Olarunpasegon,A, Pal,S, Parke,K, Pasternak,S, Paul,H, Perez,A, Perez,L, Pfannkuch,C, Plopper,F, Poindexter,A, Popovic,D, Primus,E, Pu,L, L, Puzo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M, A, Reigh,R, Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rives,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S, J, Sanders,W, Savery,G, Scherer,S, Scott,G, Shatman,S, Shen,H, Shetty,J, Shvartsbeyn,A, Sison,I, Sitter,C, D, Snaj,D, Sneed,A, Sodergren,E, Song,X, Z, Sorelle,R, Sosa,J, Steime,M, Strong,R, Sutton,A, Swalek,A, Tabot,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Usmani,K, Valas,R, Vera,V, Villaseana,D, Waldron,L, Walker,B, Wang,J, Wang,O, Wang,S, Warren,J, Warren,R, Wei,X, White,F, Williams,G, Willson,R, Wicznyk,R, Wooden,H, Morley,K, Wright,D, Wright,R, Wu,Y, Yakub,S, Yen,J, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von Niederhausen,A, Weiss,R, Smith,D, R, Holt,R, A, Smith,H, O, Weinstock,G, and Gibbs,R, A.

TITLE

Direct Submission

REFERENCE

2 (bases 1 to 239980)

Rat Genome Sequencing Consortium.

TITLE

Direct Submission

REFERENCE

Submitted (18-SEP-2002)

Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE

3 (bases 1 to 239980)

Rat Genome Sequencing Consortium.

TITLE

Direct Submission

REFERENCE

Submitted (15-NOV-2002)

Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE

On Nov 15, 2002

this sequence version replaced gi:23096650. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/project/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: KCHB

Center clone name: CH230-1116

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 209102 bases at least Q40

Consensus quality: 215667 bases at least Q20

Estimated insert size: 215679; sum-of-contigs estimation

Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/genbankdraft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 6 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1 10311: contig of 10311 bp in length

10312 10411: gap of unknown length

10412 23186: contig of 22145 bp in length

23187 23196: gap of unknown length

23197 23517: contig of 3206 bp in length

23518 23527: gap of unknown length

23528 23571: contig of 1446 bp in length

23572 23681: gap of unknown length

23682 23809: contig of 1277 bp in length

23810 23819: gap of unknown length

23820 23980: contig of 1785 bp in length.

23981 23998: contig of 1785 bp in length.

23999 23998: contig of 1785 bp in length.

24000 23998: contig of 1785 bp in length.

24001 23998: contig of 1785 bp in length.

24002 23998: contig of 1785 bp in length.

24003 23998: contig of 1785 bp in length.

24004 23998: contig of 1785 bp in length.

24005 23998: contig of 1785 bp in length.

24006 23998: contig of 1785 bp in length.

24007 23998: contig of 1785 bp in length.

24008 23998: contig of 1785 bp in length.

24009 23998: contig of 1785 bp in length.

24010 23998: contig of 1785 bp in length.

24011 23998: contig of 1785 bp in length.

24012 23998: contig of 1785 bp in length.

24013 23998: contig of 1785 bp in length.

24014 23998: contig of 1785 bp in length.

24015 23998: contig of 1785 bp in length.

24016 23998: contig of 1785 bp in length.

24017 23998: contig of 1785 bp in length.

24018 23998: contig of 1785 bp in length.

24019 23998: contig of 1785 bp in length.

24020 23998: contig of 1785 bp in length.

24021 23998: contig of 1785 bp in length.

24022 23998: contig of 1785 bp in length.

24023 23998: contig of 1785 bp in length.

24024 23998: contig of 1785 bp in length.

REFERENCE		The Chimpanzee Chromosome Y Sequencing Consortium.
AUTHORS		DNA sequence of chimpanzee chromosome Y and its evolutionary implications
TITLE		Unpublished
JOURNAL		
REFERENCES		2 (bases 1 to 171989)
AUTHORS		Hattori,M., Toyoda,A., Noguchi,H., Taylor,T.D., Kuroki,Y., Fujiyama,A. and Sakaki,Y.
JOURNAL		Direct Submission
TITLE		Submitted (28-JUN-2004) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
FEATURES		(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
SOURCE		Location/Qualifiers
		1. 171989
		/organism="Pan troglodytes"
		/mol_type="genomic DNA"
		/db_xref="taxon:9598"
		/chromosome="Y"
		/clone="PTB-386A09"
ORIGIN		
Query Match	14.0%;	Score 56; DB 5; Length 171989;
Best Local Similarity	65.6%;	Pred. No. 2.3e-07;
Matches	80; Conservative	1; Mismatches 41; Indels 0; Gaps 0;
Oy	138	TTTGATATATATATTATACATACAGACACACACACACACATATTTTAACCTCGG 197
Db	94930	TTCTATATATATATATATATATACACACACACACACACACACACACACAAATTGGG 94871
Oy	198	GACATAAATGTATACACCACCTTATCAAGTAAAAACTACTCTCAGAAGTAAGACATTA 257
Db	94870	GTTATATATCAAATPAGTCGACGTTCAAACTTATPAAGTCTATGGAAGAAAACAAGCAACA 94811
Oy	258 TC 259	
Db	94810 CC 94809	
RESULT 29		
LOCUS	AC151479	
DEFINITION	Pan troglodytes BAC clone CH251-656P8 from chromosome Y, complete sequence.	
ACCESSION	AC151479	
VERSION	AC151479.4	GI:66773775
KEYWORDS	HTG.	
SOURCE	Pan troglodytes (chimpanzee)	
ORGANISM	Pan troglodytes	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Pan.
AUTHORS		1 (bases 1 to 179557)
TITLE		Trani,L. and Haglund,K.
JOURNAL		The sequence of Pan troglodytes BAC clone CH251-656P8 Unpublished (2001)
REFERENCE		2 (bases 1 to 179557)
AUTHORS		Wilson,R.K.
TITLE		Direct Submission
JOURNAL		Submitted (02-SEP-2004) Genetics, Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE		3 (bases 1 to 179557)
AUTHORS		Wilson,R.K.
TITLE		Direct Submission
JOURNAL		Submitted (06-APR-2005) Genetics, Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE		4 (bases 1 to 179557)
AUTHORS		Wilson,R.K.
TITLE		Direct Submission
JOURNAL		Submitted (27-MAY-2005) Genetics, Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE		5 (bases 1 to 179557)

JOURNAL Submitted (09-MAY-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Nov 8, 2000 this sequence version replaced gi:9331947.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0109G18

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
The position of this clone was established as part of a collaboration between the Human Chromosome Y Mapping Project (Tomoko Kawaguchi, Helen Skalecky, Laura G. Brown, Steve Rozen, and David C. Page at the Whitehead Institute for Biomedical Research, Cambridge MA) and the Washington University Genome Sequencing Center, St. Louis MO.

SOURCE INFORMATION:
The RPL1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tatem, M., Calaneese, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.research.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)
VECTOR: pBACE3.6
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RPL1-95B23, 200 bp overlap; the clone sequenced to the right is RPL1-477B5, 200 bp overlap. Actual start of this clone is at base position 195 of RPL1-95B23; actual end is at base position 54795 of RPL1-477B5.

FEATURES

source
1..60810
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="Y"
 /map="Y"
 /clone="RPL1-109G18"
 /clone_1lb="RPL1-11"
repeat_region
 /rpc_family="Alu"
 265..560
repeat_region
 /rpc_family="Alu"
 2076..2380
repeat_region
 /rpc_family="L1"
 2450..2738
repeat_region
 /rpc_family="L1"
 2745..2803
repeat_region
 /rpc_family="Alu"
 3684..3957
repeat_region
 /rpc_family="Alu"
 4467..4758
repeat_region
 /rpc_family="Alu"
 5026..5165
repeat_region
 /rpc_family="MALR"

repeat_region 5167..5331
 /rpc_family="Alu"
repeat_region 5332..5494
 /rpc_family="L1"
repeat_region 5495..5624
 /rpc_family="Alu"
repeat_region 5625..5773
 /rpc_family="MALR"
repeat_region 5777..6089
 /rpc_family="ERV1"
repeat_region 6187..6275
 /rpc_family="ERV1"
repeat_region 6488..6734
 /rpc_family="Alu"
repeat_region 6735..7039
 /rpc_family="ERV1"
repeat_region 7084..7256
 /rpc_family="Alu"
repeat_region 7792..7840
 /rpc_family="ERV1"
repeat_region 11645..11947
 /rpc_family="Alu"
repeat_region 12684..12930
 /rpc_family="Alu"
repeat_region 13098..13393
 /rpc_family="Alu"
repeat_region 15135..15248
 /rpc_family="Alu"
repeat_region 15249..15704
 /rpc_family="ERV1"
repeat_region 15810..16105
 /rpc_family="Alu"
repeat_region 16168..17185
 /rpc_family="ERV1"
repeat_region 18582..18676
 /rpc_family="ERV1"
repeat_region 19702..19790
 /rpc_family="L1"
repeat_region 19822..19889
 /rpc_family="L1"
repeat_region 20026..20093
 /rpc_family="L1"
repeat_region 20116..20168
 /rpc_family="L1"
repeat_region 20456..21092
 /rpc_family="L1"
repeat_region 21147..21588
 /rpc_family="L1"
repeat_region 21629..21913
 /rpc_family="Alu"
repeat_region 21944..22564
 /rpc_family="L1"
repeat_region 22369..22679
 /rpc_family="L1"
repeat_region 22696..23103
 /rpc_family="L1"
repeat_region 23362..23512
 /rpc_family="L1"
repeat_region 23851..24119
 /rpc_family="Alu"
repeat_region 24347..24425
 /rpc_family="ERV"
repeat_region 24466..24973
 /rpc_family="ERV"
repeat_region 28368..29123
 /rpc_family="ERV1"
repeat_region 29132..29333
 /rpc_family="ERV1"
repeat_region 30293..30562
 /rpc_family="ERV"
repeat_region 30570..30920
 /rpc_family="MALR"
repeat_region 31587..32003

REFERENCE 3 (bases 1 to 212207)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (10-SEP-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 212207)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (11-OCT-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 5 (bases 1 to 212207)
 AUTHORS Waterston,R.
 TITLE Direct Submission
 JOURNAL Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA On Oct 11, 2001 this sequence version replaced gi:1578784.

----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@wustl.wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0502M01
 Drafting Center: WIBR

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateo,M., Catalanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.reagen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the right is RP11-219G14. Actual start of this clone is at base position 1 of RP11-502M1; actual end is at base position 212207 of RP11-502M1.

The sequence of AC013337 has been incorporated into AC093853.

FEATURES
 source
 Location/Qualifiers
 1..212207
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="4"
 /map="4"
 /clone="RP11-502M1"
 /clone_id="RPCI-11"
 repeat_region
 239..408

repeat_region
 /rpt_family="L2"
 410..536
 /rpt_family="L2"
 repeat_region
 891..1138
 /rpt_family="L2"
 repeat_region
 1658..1694
 /rpt_family="AT_rich"
 1695..1768
 /rpt_family="TA)n"
 2079..2362
 /rpt_family="TA)n"
 2377..2616
 /rpt_family="TA)n"
 4159..4293
 /rpt_family="MIR"
 4906..5042
 /rpt_family="MER1_type"
 5740..6146
 /rpt_family="L2"
 6853..6981
 /rpt_family="TA)n"
 7233..7524
 /rpt_family="Alu"
 8043..8532
 /rpt_family="TA)n"
 8652..8724
 /rpt_family="ERV1"
 8666..8722
 /rpt_family="CA)n"
 8770..8946
 /rpt_family="L1"
 9683..9712
 /rpt_family="AT_rich"
 9813..9837
 /rpt_family="AT_rich"
 12004..12524
 /rpt_family="MER1_type?"
 12505..12558
 /rpt_family="AT_rich"
 12594..12615
 /rpt_family="MER1_type?"
 12615..12645
 /rpt_family="(T)n"
 12616..12926
 /rpt_family="Alu"
 12927..13003
 /rpt_family="MER1_type?"
 13262..13631
 /rpt_family="MAIR"
 13632..15201
 /rpt_family="MAIR"
 15202..15564
 /rpt_family="MAIR"
 15700..15746
 /rpt_family="L2"
 15858..16368
 /rpt_family="BRV1"
 16461..17334
 /rpt_family="ERVX"
 17348..17734
 /rpt_family="ERVX"
 17901..17928
 /rpt_family="(TG)n"
 19062..19133
 /rpt_family="AT_rich"
 19172..19193
 /rpt_family="AT_rich"
 19928..20049
 /rpt_family="L2"
 20068..20477
 /rpt_family="L1"
 20832..21061
 /rpt_family="L1"

Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dethorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaey, K.R., Delgado, O., Dunn, A.L., Ding, Y., Dinh, H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., He, X., Hernandez, J., Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B., Homsy, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolyet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovach, C., Kratovic, J., Kureish, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, Z., Li, Z., Lichtarge, O., Lien, C., Liu, J., Liu, W., Louisedg, H., Lozdo, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martine, E., Massey, E., Maxwell, E., McLeod, M.P., Meador, M., Mel, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabhat, K., Moore, S., Morgan, M., Moorish, T., Morris, S., Moser, M., Neal, D., Nelson, D., Newsum, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguni, M., Okunonu, G., Otagunye, N., Oyedero, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M., Rojaf, A., Rojchokan, I., Rolle, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshitrali, N., Sisson, I., Sodergren, E., Sonake, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Teifrod, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalton, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wlarczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Zyllator, S.L., Westminster, G., and Gibbs, R.

JOURNAL	Unpublished
REFERENCE	2 (pages 1 to 176108)
AUTHORS	Worley, K.C.

JOURNAL Submitted (19-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

AUTHORS Worley, K.C.
TITLE Direct Submission
JOURNAL Submitted (29-APR-2002) Human Genome Sequencing Center, Department

REFERENCE Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 176108)
AUTHORS Worley, K.C.

JOURNAL Submitted (30-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

INFORMATION: <http://www.ngsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones.

Features listing.

ANNOTATION OF FEATURES:

of a local database that includes entries from dbSTS, GDB, and local mapping efforts. Repeats are identified using RepeatMasker (A. Smit and P. Green unpublished.) for Human and Mouse sequences. Genes and Region of sequence similarity are identified by BLAST

```
repeat_region      /xpt_family="AT_rich"
                   14756. .14963
```

REFERENCE	1 (bases 1 to 184182)	Stanford Human Genome Center and Los
AUTHORS	DOE Joint Genome Institute,	
	Alamos National Laboratory.	
TITLE	Direct Submission	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 184182)	
AUTHORS	DOE Joint Genome Institute.	

REFERENCE	TITLE	GENOME	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	TITLE	GENOME	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM
REFERENCE	JOURNAL	Submitted (31-AUG-2001)	Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA	3 (bases 1 to 184182)	DOE Joint Genome Institute.	Direct Submission			REFERENCE	JOURNAL	Submitted (18-MAR-2003)	DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA	On Mar 18, 2003 this sequence version replaced gi:17298600.	Direct Sequence Produced by DOE Joint Genome Institute	www.jgi.doe.gov	Fishing Completed at Stanford Human Genome Center and Los Alamos National Laboratory	www-sngc.stanford.edu
REFERENCE	JOURNAL	Submitted (04-DEC-2001)	Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA	4 (bases 1 to 184182)	DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.	Direct Submission			REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice
REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice	Dept. Genomics and Bioinformatics	Roslin Institute, Roslin							
REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice	Dept. Genomics and Bioinformatics	Roslin Institute, Roslin							
REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice	Dept. Genomics and Bioinformatics	Roslin Institute, Roslin							
REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice	Dept. Genomics and Bioinformatics	Roslin Institute, Roslin							
REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice	Dept. Genomics and Bioinformatics	Roslin Institute, Roslin							
REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice	Dept. Genomics and Bioinformatics	Roslin Institute, Roslin							
REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice	Dept. Genomics and Bioinformatics	Roslin Institute, Roslin							
REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice	Dept. Genomics and Bioinformatics	Roslin Institute, Roslin							
REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice	Dept. Genomics and Bioinformatics	Roslin Institute, Roslin							
REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice	Dept. Genomics and Bioinformatics	Roslin Institute, Roslin							
REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice	Dept. Genomics and Bioinformatics	Roslin Institute, Roslin							
REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice	Dept. Genomics and Bioinformatics	Roslin Institute, Roslin							
REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice	Dept. Genomics and Bioinformatics	Roslin Institute, Roslin							
REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice	Dept. Genomics and Bioinformatics	Roslin Institute, Roslin							
REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice	Dept. Genomics and Bioinformatics	Roslin Institute, Roslin							
REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice	Dept. Genomics and Bioinformatics	Roslin Institute, Roslin							
REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice	Dept. Genomics and Bioinformatics	Roslin Institute, Roslin							
REFERENCE	JOURNAL	Submitted (16-JUL-2001)	Morrice D., Roslin Institute, Roslin, Midlothian, EH25 9PS, UK	David Morrice@dbbrc.ac.uk	Vector pBLUESCRIPT SK+	R. Site 1 Ecot1	R. Site 2 Ecot1	Contact: David Morrice	Dept. Genomics and Bioinformatics	Roslin Institute, Roslin							

<p>FEATURES</p> <p>source</p> <p>STS</p> <p>ORIGIN</p>	<p>Midlochian, EH25 9PS, UK Tel: +44 (0) 131 527 4200 Fax: +44 (0) 131 440 0434 Email: david.morrice@bsrc.ac.uk Seq primer T3.</p> <p>Location/Qualifiers</p> <pre> 1..668 /organism="Meleagris gallopavo" /mol_type="genomic DNA" /db_xref="taxon:9103" 1..668 /standard_name="TC1458"</pre>	<p>Query Match</p> <p>Best Local Similarity 58.9%; Pred.No. 1.9e-07; Matches 93; Conservative 1; Mismatches 64; Indels 0; Gaps 0;</p>	<p>QY</p> <p>129 TTGTGACTGTTGATATATATATATATACACACACACACACACACATATT 188 217 TTCTTATTCGTATATATATACATATACACACACACACACACACACATATA 158 </p> <p>Db</p> <p>189 TTAACTCGGGACAPTAATAGTGACACCACCATTCGAAGTAAAACATATCACAAGA 248 157 TAATGCATATGTTTACCATATTCCTAGGAATTTATACACATCAAGAAAAATCGACAGTC 98 </p> <p>QY</p> <p>249 AAGACATATATCTGCTATTTGTCTATTACTAGACAGTTAA 286 97 TAGTGAAAAATCTCGACACACTGATCATGAAACTGACAA 60 </p>	<p>RESULT 38</p> <p>AL593618 668 bp DNA linear STS 16-JUL-2001</p> <p>LOCUS Meleagris gallopavo DNA; STS TC1458, sequence tagged site. DEFINITION ACCESSION AL593618 VERSION AL593618.1 GI:14716736 KEYWORDS SOURCE STS. ORGANISM Meleagris gallopavo (turkey) Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Meleagris.</p>	<p>REFERENCE</p> <p>AUTHORS Morrice,D. JOURNAL Unpublished REFERENCE 2 (bases 1 to 668)</p> <p>AUTHORS Morrice,D. TITLE Direct Submission JOURNAL Submitted (16-JUL-2001) Morrice D., Roslin Institute, Roslin, MIDLOCHIAN, EH25 9PS, UK david.morrice@bsrc.ac.uk COMMENT Vector pBLUESCRIPT SK+ R. Site 1 EcorI R. Site 2 EcorI Contact: David Morrice Dept. Genomics and Bioinformatics Roslin Institute, Roslin Midlochian, EH25 9PS, UK Tel: +44 (0) 131 527 4200 Fax: +44 (0) 131 440 0434 Email: david.morrice@bsrc.ac.uk Seq primer T7.</p>	<p>FEATURES</p> <p>source</p> <p>STS</p> <p>ORIGIN</p>	<p>1..668 /organism="Meleagris gallopavo" /mol_type="genomic DNA" /db_xref="taxon:9103" 1..668 /standard_name="TC1458"</p>	<p>Query Match</p> <p>Best Local Similarity 58.9%; Score 55.2; DB 7; Length 668; Matches 93; Conservative 1; Mismatches 64; Indels 0; Gaps 0;</p>	<p>QY</p> <p>129 TTGTGACTGTTGATATATATATATATACACACACACACACACACATATT 188 217 TTCTTATTCGTATATATATACATATACACACACACACACACACACATATA 158 </p> <p>Db</p> <p>189 TTAACTCGGGACAPTAATAGTGACACCACCATTCGAAGTAAAACATATCACAAGA 248 157 TAATGCATATGTTTACCATATTCCTAGGAATTTATACACATCAAGAAAAATCGACAGTC 98 </p> <p>QY</p> <p>249 AAGACATATATCTGCTATTTGTCTATTACTAGACAGTTAA 286 97 TAGTGAAAAATCTCGACACACTGATCATGAAACTGACAA 60 </p>	<p>RESULT 38</p> <p>AL593618 668 bp DNA linear STS 16-JUL-2001</p> <p>LOCUS Meleagris gallopavo DNA; STS TC1458, sequence tagged site. DEFINITION ACCESSION AL593618 VERSION AL593618.1 GI:14716736 KEYWORDS SOURCE STS. ORGANISM Meleagris gallopavo (turkey) Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Meleagris.</p>	<p>REFERENCE</p> <p>AUTHORS Morrice,D. JOURNAL Unpublished REFERENCE 2 (bases 1 to 668)</p> <p>AUTHORS Morrice,D. TITLE Direct Submission JOURNAL Submitted (16-JUL-2001) Morrice D., Roslin Institute, Roslin, MIDLOCHIAN, EH25 9PS, UK david.morrice@bsrc.ac.uk COMMENT Vector pBLUESCRIPT SK+ R. Site 1 EcorI R. Site 2 EcorI Contact: David Morrice Dept. Genomics and Bioinformatics Roslin Institute, Roslin Midlochian, EH25 9PS, UK Tel: +44 (0) 131 527 4200 Fax: +44 (0) 131 440 0434 Email: david.morrice@bsrc.ac.uk Seq primer T7.</p>	<p>FEATURES</p> <p>source</p> <p>STS</p> <p>ORIGIN</p>	<p>1..668 /organism="Meleagris gallopavo" /mol_type="genomic DNA" /db_xref="taxon:9103" 1..668 /standard_name="TC1458"</p>	<p>Query Match</p> <p>Best Local Similarity 58.9%; Score 55.2; DB 7; Length 668; Matches 93; Conservative 1; Mismatches 64; Indels 0; Gaps 0;</p>	<p>QY</p> <p>129 TTGTGACTGTTGATATATATATATATACACACACACACACACACATATT 188 217 TTCTTATTCGTATATATATACATATACACACACACACACACACACATATA 158 </p> <p>Db</p> <p>189 TTAACTCGGGACAPTAATAGTGACACCACCATTCGAAGTAAAACATATCACAAGA 248 157 TAATGCATATGTTTACCATATTCCTAGGAATTTATACACATCAAGAAAAATCGACAGTC 98 </p> <p>QY</p> <p>249 AAGACATATATCTGCTATTTGTCTATTACTAGACAGTTAA 286 97 TAGTGAAAAATCTCGACACACTGATCATGAAACTGACAA 60 </p>	<p>RESULT 38</p> <p>AL593618 668 bp DNA linear STS 16-JUL-2001</p> <p>LOCUS Meleagris gallopavo DNA; STS TC1458, sequence tagged site. DEFINITION ACCESSION AL593618 VERSION AL593618.1 GI:14716736 KEYWORDS SOURCE STS. ORGANISM Meleagris gallopavo (turkey) Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Meleagris.</p>	<p>REFERENCE</p> <p>AUTHORS Morrice,D. JOURNAL Unpublished REFERENCE 2 (bases 1 to 668)</p> <p>AUTHORS Morrice,D. TITLE Direct Submission JOURNAL Submitted (16-JUL-2001) Morrice D., Roslin Institute, Roslin, MIDLOCHIAN, EH25 9PS, UK david.morrice@bsrc.ac.uk COMMENT Vector pBLUESCRIPT SK+ R. Site 1 EcorI R. Site 2 EcorI Contact: David Morrice Dept. Genomics and Bioinformatics Roslin Institute, Roslin Midlochian, EH25 9PS, UK Tel: +44 (0) 131 527 4200 Fax: +44 (0) 131 440 0434 Email: david.morrice@bsrc.ac.uk Seq primer T7.</p>	<p>FEATURES</p> <p>source</p> <p>STS</p> <p>ORIGIN</p>	<p>1..668 /organism="Meleagris gallopavo" /mol_type="genomic DNA" /db_xref="taxon:9103" 1..668 /standard_name="TC1458"</p>	<p>Query Match</p> <p>Best Local Similarity 58.9%; Score 55.2; DB 7; Length 668; Matches 93; Conservative 1; Mismatches 64; Indels 0; Gaps 0;</p>	<p>QY</p> <p>129 TTGTGACTGTTGATATATATATATATACACACACACACACACACATATT 188 217 TTCTTATTCGTATATATATACATATACACACACACACACACACACATATA 158 </p> <p>Db</p> <p>189 TTAACTCGGGACAPTAATAGTGACACCACCATTCGAAGTAAAACATATCACAAGA 248 157 TAATGCATATGTTTACCATATTCCTAGGAATTTATACACATCAAGAAAAATCGACAGTC 98 </p> <p>QY</p> <p>249 AAGACATATATCTGCTATTTGTCTATTACTAGACAGTTAA 286 97 TAGTGAAAAATCTCGACACACTGATCATGAAACTGACAA 60 </p>	<p>RESULT 38</p> <p>AL593618 668 bp DNA linear STS 16-JUL-2001</p> <p>LOCUS Meleagris gallopavo DNA; STS TC1458, sequence tagged site. DEFINITION ACCESSION AL593618 VERSION AL593618.1 GI:14716736 KEYWORDS SOURCE STS. ORGANISM Meleagris gallopavo (turkey) Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Meleagris.</p>	<p>REFERENCE</p> <p>AUTHORS Morrice,D. JOURNAL Unpublished REFERENCE 2 (bases 1 to 668)</p> <p>AUTHORS Morrice,D. TITLE Direct Submission JOURNAL Submitted (16-JUL-2001) Morrice D., Roslin Institute, Roslin, MIDLOCHIAN, EH25 9PS, UK david.morrice@bsrc.ac.uk COMMENT Vector pBLUESCRIPT SK+ R. Site 1 EcorI R. Site 2 EcorI Contact: David Morrice Dept. Genomics and Bioinformatics Roslin Institute, Roslin Midlochian,</p>
---	---	--	--	--	--	---	--	--	--	--	--	---	--	--	--	--	--	---	--	--	--	--	--


```

* 27258 27357: gap of 100 bp
* 27358 28067: contig of 720 bp in length
* 28068 28167: gap of 100 bp
* 28168 28876: contig of 709 bp in length
* 28877 28977: gap of 100 bp
* 28978 29681: contig of 705 bp in length
* 29682 29781: gap of 100 bp
* 29782 30488: contig of 707 bp in length
* 30489 30588: gap of 100 bp
* 30589 31289: contig of 701 bp in length
* 31290 31390: gap of 100 bp
* 31391 32112: contig of 723 bp in length
* 32113 32312: gap of 100 bp
* 32313 32911: contig of 699 bp in length
* 32912 33011: gap of 100 bp
* 33012 33724: contig of 713 bp in length
* 33725 33824: gap of 100 bp
* 33825 34521: contig of 697 bp in length
* 34522 34621: gap of 100 bp
* 34622 35321: contig of 700 bp in length
* 35322 35421: gap of 100 bp
* 35422 36128: contig of 707 bp in length
* 36129 36228: gap of 100 bp
* 36229 36945: contig of 717 bp in length
* 36946 37045: gap of 100 bp
* 37046 37757: contig of 712 bp in length
* 37758 37857: gap of 100 bp
* 37858 38559: contig of 702 bp in length
* 38560 38659: gap of 100 bp
* 38660 39351: contig of 692 bp in length
* 39352 39451: gap of 100 bp
* 39452 40159: contig of 708 bp in length
* 40160 40259: gap of 100 bp
* 40260 40956: contig of 697 bp in length
* 40957 41056: gap of 100 bp
* 41057 41776: contig of 720 bp in length
* 41777 41876: gap of 100 bp
* 41877 42592: contig of 716 bp in length
* 42593 42692: gap of 100 bp
* 42693 43397: contig of 705 bp in length
* 43398 43497: gap of 100 bp
* 43498 44195: contig of 698 bp in length
* 44196 44295: gap of 100 bp
* 44296 45030: contig of 735 bp in length
* 45031 45130: gap of 100 bp
* 45131 45841: contig of 711 bp in length
* 45842 45941: gap of 100 bp
* 45942 46660: contig of 719 bp in length
* 46661 46761: gap of 100 bp
* 46762 47473: contig of 713 bp in length
* 47474 47573: gap of 100 bp
* 47574 48287: contig of 714 bp in length
* 48288 48387: gap of 100 bp
* 48388 49102: contig of 715 bp in length
* 49103 49202: gap of 100 bp
* 49203 49808: contig of 706 bp in length
* 49809 50008: gap of 100 bp
* 50009 50736: contig of 728 bp in length
* 50737 50836: gap of 100 bp
* 50837 51527: contig of 691 bp in length
* 51528 51627: gap of 100 bp
* 51628 52327: contig of 700 bp in length
* 52328 52427: gap of 100 bp
* 52428 53145: contig of 718 bp in length
* 53146 53245: gap of 100 bp
* 53246 53974: contig of 729 bp in length
* 53975 54074: gap of 100 bp
* 54075 54789: contig of 715 bp in length
* 54790 54889: gap of 100 bp
* 54890 55586: contig of 697 bp in length

```

Query Match
Best Local Similarity

13.8% Score 55.2; DB 12; Length 61255;
66.1% Pred. No. 3.7e-07;

```

Matches 78; Conservative 1; Mismatches 39; Indels 0; Gaps 0;
QY 119 GACCTACCTTTTGTACTGTTGATATATATATATACACACACACACACACA 178
Db 9818 GGTCTACGATACATGTAACAAATACACACACACACACACACACACACACA 9759
QY 179 CACATATATTTTAACTCGGACATTAATGTCACACACCTATTCAAGTAAACTACT 236
Db 9758 CACACAAACATTTCATCGCACATTAATATGCTACATATATGCAACATACATACCT 9701

RESULT 40
AL590437
LOCUS AL590437 151358 bp DNA linear HTG 10-JUL-2001
DEFINITION Homo sapiens chromosome 1 clone RP11-394E12, 6 unordered pieces.
ACCESSION AL590437
VERSION AL590437.4 GI:13620473
KEYWORDS HTG; HTGS PHASE1; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1
REFERENCE
AUTHORS McLay, K.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT On Apr 12, 2001 this sequence version replaced gi:13619301.

----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: ba394E12
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator ET-amersham; 4% of reads
Dye-terminator Big Dye; 95% of reads
Consensus quality: 149170 bases at least Q40
Consensus quality: 149671 bases at least Q30
Consensus quality: 150169 bases at least Q20
Insert size: 148776; 6.0% error; agarose-fp
Quality coverage: 7.08x in Q20 bases; sum-of-contigs
Quality coverage: 7.18x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 22354: contig of 22354 bp in length
* 22355 22454: gap of 100 bp
* 22455 27105: contig of 4651 bp in length
* 27106 27205: gap of 100 bp
* 27206 47237: contig of 20032 bp in length
* 47238 47337: gap of 100 bp
* 47338 64870: contig of 17533 bp in length
* 64871 64970: gap of 100 bp
* 64971 74039: contig of 9068 bp in length
* 74039 74138: gap of 100 bp
* 74139 151358: contig of 77220 bp in length.
FEATURES
source
1..151358
/organism="Homo sapiens"
/mol_type="genomic DNA"

```


repeat_region	3579..3637	/rpt_family="Alu"
repeat_region	3663..3848	/rpt_family="L2"
repeat_region	4553..4654	/rpt_family="L2"
repeat_region	4651..4755	/rpt_family="L2"
repeat_region	5631..8264	/rpt_family="L1"
repeat_region	9547..9852	/rpt_family="MER1_type"
repeat_region	9990..10054	/rpt_family="MER1_type"
repeat_region	10252..10282	/rpt_family="AT_rich"
repeat_region	10279..10478	/rpt_family="MER1_type"
repeat_region	10615..10640	/rpt_family="AT_rich"
repeat_region	10671..10963	/rpt_family="Alu"
repeat_region	10957..10990	/rpt_family="(A)n"
repeat_region	12173..12277	/rpt_family="TC2"
repeat_region	13689..13997	/rpt_family="L1"
repeat_region	14211..14412	/rpt_family="L2"
repeat_region	14411..14729	/rpt_family="L2"
repeat_region	14777..15003	/rpt_family="MR"
repeat_region	15039..15403	/rpt_family="MER2_type"
repeat_region	15400..15510	/rpt_family="(PATATG)n"
repeat_region	15506..15629	/rpt_family="L2"
repeat_region	15630..15830	/rpt_family="MR"
repeat_region	15866..15979	/rpt_family="L2"
repeat_region	16071..16491	/rpt_family="MALR"
repeat_region	17261..17507	/rpt_family="Mariner"
repeat_region	18707..19811	/rpt_family="L1"
repeat_region	19686..19717	/rpt_family="AT_rich"
repeat_region	19812..20111	/rpt_family="Alu"
repeat_region	20112..20956	/rpt_family="L1"
repeat_region	20965..21444	/rpt_family="L1"
repeat_region	21463..21522	/rpt_family="AT_rich"
repeat_region	21567..21816	/rpt_family="L1"
repeat_region	21890..22003	/rpt_family="MER1_type"
repeat_region	22037..22135	/rpt_family="MER1_type"
repeat_region	22576..22615	/rpt_family="(T)n"
repeat_region	22650..22668	/rpt_family="(TA)n"
repeat_region	22686..22715	/rpt_family="(CA)n"
repeat_region	29023..29196	

```

repeat_region      /rpt_family="MER1_type"
                    29122..29151
misc_feature        /rpt_family="AT_rich"
                    31344..31706
                    /note="similar to EST BF512614 (NID:g11597793) "
repeat_region      31641..33029
                    /rpt_family="L2"
repeat_region      32184..32382
                    /rpt_family="L2"
repeat_region      32403..32543
                    /rpt_family="CR1"
repeat_region      32562..32781
                    /rpt_family="MIR"
repeat_region      32803..32949
                    /rpt_family="CR1"
misc_feature        32931..33395
                    /note="similar to EST AW236561 (NID:g6566950) xm47g11.x1"
repeat_region      34086..34219
                    /rpt_family="MIR"
misc_feature        34117..34473
                    /note="similar to EST AA889061 (NID:g3015940) am34h02.s1"
repeat_region      34965..37833
                    /rpt_family="L1"
repeat_region      37854..38167
                    /rpt_family="ALU"
repeat_region      38704..38737

```

Query Match	13.8%	Score 55.2;	DB 5;	Length 190080;
Best Local Similarity	70.6%	Pred. NO. 4.4e-07;		
Matches	72;	Conservative	1;	Mismatches 29;
			Indels	0;
			Gaps	0

[illegible]

RESULT	42			
AC136742/c				
LOCUS		197131 bp	DNA	linear
DEFINITION	Mus musculus	chromosome 5, clone RP23-145H8,	complete sequence.	
ACCESSION	AC136742			
VERSION	AC136742.15	GI:87116735		
KEYWORDS	HTG.			
SOURCE	Mus musculus	(house mouse)		
ORGANISM	Mus musculus			

REFERENCE
AUTHORS
JOURNAL
TITLE
REFERENCE

2 (bases 1 to 197131)
Mus musculus chromosome 5, clone RP23-145H8
Unpublished
2 (bases 1 to 197131)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bonkshgiller, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gird, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagoe, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kellis, C., Lander, S., Levine, R.,
Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J.,
Matthews, C., McCarthy, M., Meldrum, J., Menais, L., Mohova, T.,
Miyata, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Plunkhang, P., Pierre, N., Raymond, C., Retta, R.,
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnuppach, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Talamas, J., Testafte, S., Theodore, J., Topham, K.,

Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. Submitted (07-NOV-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
JOURNAL
TITLE
SUBMITTED (10-FEB-2006) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS
Birten, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArrellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., McLean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Menga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramaamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
SUBMITTED (14-APR-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS
Birten, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArrellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., McLean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Menga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramaamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
SUBMITTED (25-JUN-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS
Birten, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArrellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., McLean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Menga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramaamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
SUBMITTED (10-FEB-2006) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Broad Institute of MIT and Harvard
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
Project Information
Center project name: L26537
Center clone name: 145_H_8

FEATURES
source
1.197133
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="5"
/map="5"
/clone="RP23-145H8"
/clone_lib="RP23 Female Mouse BAC"
1.56
/note="PCR product sequence only"
902.936
/rpt_family="TG)n"
complement(937..1276)
/rpt_family="RLR10B2"
1288.1440
/rpt_family="TG)n"
1459.1547
/rpt_family="TG)n"
1510.1569
/note="single clone coverage"
1573.1664
/rpt_family="TG)n"
1707.1886
/rpt_family="TG)n"
2455.2596
/rpt_family="WT2a"
2510.2714
/rpt_family="WT2"
complement(2715..3347)
/rpt_family="RMR19C"
4986.5202
/rpt_family="ORR1E"
5225.5546
/rpt_family="WT2"
5954.6028
/rpt_family="Lx8"
6150.6325
/rpt_family="Lx8"
6408.6431
/rpt_family="AT_rich"
complement(7483..7653)
/rpt_family="Lx8"
7927.7975
/rpt_family="CA)n"
8180.8271
/rpt_family="MER58A"
complement(9758..10258)
/rpt_family="RMR15"
10708.10885
/rpt_family="TCRA)n"
10914.10954
/rpt_family="CAAA)n"

Muzylo, D., Warte, E., Metzker, M., Lee, S., Abramzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, E., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Bacc, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barneshead, M., Bernabed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Buttsell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, E., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., L., Davis, C., Day, Carroll, L., De Anda, C., Dedetich, D., Delgado, O., Denson, S., Detamco, C., Ding, Y., Dihn, H., Divya, K., Draper, H., Dugan-Roches, S., Dunn, A., Dublin, K., Duval, B., Eaves, K., Egan, A., Escoto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, P., Fraser, C. M., Gabitis, A., Gante, R., Garcia, A., Garner, T., Garza, M., Garbaterre, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guetara, W., Gunaratne, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jacob, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.

COMMENT

```

Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

-----
Project Information
Center project name: GHY
Center clone name: CH230-133C19

-----
Summary Statistics
Assembly program: Atlas 3.0.1
Consensus quality: 21393 bases at least Q40
Consensus quality: 21747 bases at least Q30
Consensus quality: 220158 bases at least Q20
Estimated insert size: 228571, sum-of-contigs estimation
Quality coverage: 6x in Q20 bases, sum-of-contigs estimation

```

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently

+	314959	315058: gap of 100 bp	+
+	315059	318574: contig of 3516 bp in Jarrowh	+

JOURNAL Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On May 9, 2003 this sequence version replaced gi:22951950. The sequence in this assembly is a combination of BAC based reads

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

On nucleic - nucleic search, using sw model

Run on: July 17, 2006, 16:49:59 ; Search time 3964 Seconds
(without alignment)
703.557 Million cell updates/sec

Title: SEQ1-33670G
Perfect score: 399.2
Sequence: 1 aagaatgcctctctctcata.....taactcagatcagaacgag 400

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues
Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : N Geneseq 8: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002as: *
7: Geneseqn2002bs: *
8: Geneseqn2003as: *
9: Geneseqn2003bs: *
10: Geneseqn2003cs: *
11: Geneseqn2003ds: *
12: Geneseqn2004as: *
13: Geneseqn2004bs: *
14: Geneseqn2005s: *
15: Geneseqn2006s: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB ID	Description
C	1	397.6	99.6	83400	12	ADP079906 Human, RAD
C	2	397.6	99.6	193363	14	AED89392 Human bre
C	3	60.8	15.2	265703	14	AEE04792 Cancer-as
C	4	56	14.0	463	4	AAI92919 Human pol
C	5	56	14.0	463	9	ACH24278 Human adu
C	6	53.6	13.4	141463	11	ACN43862 Human met
C	7	52.4	13.1	3455	5	AAH41011 Human met
C	8	51.6	12.9	327	3	ADP57118 Human met
C	9	51.2	12.8	49914	13	ABD33262 Murine ca
C	10	51	12.8	3462	5	AAH41010 Human met
C	11	51	12.8	3462	5	AAH41006 Human met
C	12	51	12.8	3467	5	AAH41005 Human met
C	13	51	12.8	3467	5	AAH41009 Human met
C	14	51	12.8	3469	5	AAH41007 Human met
C	15	51	12.8	3470	5	AAH41008 Human met
C	16	51	12.8	3473	5	AAH41004 Human met
C	17	51	12.8	24000	3	AAH88551 Human dua
C	18	51	12.8	28854	8	ABX95685 Human gen

C	19	50.8	12.7	92726	9	ADA02927 Mouse pik
C	20	50.8	12.7	92726	10	ADB72665 Mouse pik
C	21	50.8	12.7	92726	10	ADC85407 Mouse pik
C	22	50.8	12.7	92726	12	ADM74522 Murine pik
C	23	50.2	12.6	101169	12	ADQ97584 Mouse can
C	24	50	12.5	5289	4	AAK67814 Human imm
C	25	50	12.5	61876	11	ACN44732 Mouse gen
C	26	50	12.5	68960	14	ADZ13128 Human met
C	27	50	12.5	110000	11	ACN43998-4
C	28	50	12.5	191127	10	ADH13788
C	29	49.8	12.5	2534	3	AAK69141 Human ABC
C	30	49.8	12.5	17970	6	ABG67033 Human ang
C	31	49.8	12.5	28829	13	ABD33193 Murine ca
C	32	49.8	12.5	159865	12	ADQ97056 Human can
C	33	49.8	12.5	188056	15	AEF19127 Human NIP
C	34	49.4	12.4	590	13	ADQ57023 Novel can
C	35	49.4	12.4	4858	6	ABK33962 Human DNA
C	36	49.4	12.4	4858	8	ADA20413 Prostate
C	37	49.4	12.4	4858	8	ADA84220 Human ren
C	38	49.4	12.4	110000	10	ADE11169-1
C	39	49.4	12.4	154394	12	ADQ17784 Human sof
C	40	49.4	12.4	154394	14	ADX06954 Cyclin-de
C	41	49.4	12.4	154394	14	ADZ80338 Human nuc
C	42	49.4	12.4	154394	15	AEF74999 Human pol
C	43	49.2	12.3	13427	6	ABL33926 Human imm
C	44	49	12.3	109565	13	ABD33086 Murine ca
C	45	49	12.3	138627	12	ADQ97183 Human can
C	46	48.8	12.2	10138	6	ABK28343 DNA trans
C	47	48.8	12.2	40633	11	ACN43848 Mouse gen
C	48	48.8	12.2	110000	13	ABD32966-01
C	49	48.8	12.2	219352	13	ABD33098 Continuation (2 of
C	50	48.8	12.2	247509	10	ADL13684 Osteoarth
C	51	48.8	12.2	247509	12	ADQ59419 Human can
C	52	48.8	12.2	247654	14	ADZ13712 Murine ca
C	53	48.6	12.2	350	4	AAI80148 Human pol
C	54	48.6	12.2	2316	6	ADY17693 DNA encod
C	55	48.6	12.2	110000	6	ABO74964-1
C	56	48.6	12.2	156843	11	ACN44786 Continuation (2 of
C	57	48.4	12.1	3205	13	ADQ72296 Human gen
C	58	48.4	12.1	18401	11	ADP75183 Human DNA
C	59	48.4	12.1	19236	6	ABN80227 Human ADA
C	60	48.4	12.1	19236	6	ABN80227 Human che
C	61	48.4	12.1	20099	10	ABZ22599 Human F13
C	62	48.4	12.1	54108	6	ABK22782 Human hig
C	63	48.4	12.1	57273	6	ABK22782 Human hig
C	64	48.4	12.1	66933	5	ABK22782 Human hig
C	65	48.4	12.1	66933	8	ABK22782 Human hig
C	66	48.4	12.1	66933	10	ABK22782 Human hig
C	67	48.4	12.1	66933	13	ABK22782 Human hig
C	68	48.4	12.1	66933	13	ABK22782 Human hig
C	69	48.4	12.1	66933	13	ABK22782 Human hig
C	70	48.4	12.1	66933	14	ABK22782 Human hig
C	71	48.4	12.1	72049	5	ABA82623 Human HBM
C	72	48.4	12.1	72049	8	ACC45364 Human HBM
C	73	48.4	12.1	72049	10	ADP88064 HBM-relat
C	74	48.4	12.1	72049	10	ADP88064 HBM-relat
C	75	48.4	12.1	72049	13	ADP88064 HBM-relat
C	76	48.4	12.1	72049	13	ADP88064 HBM-relat
C	77	48.4	12.1	72049	14	ADP88064 HBM-relat
C	78	48.4	12.1	110000	14	ADP88064 HBM-relat
C	79	48.4	12.1	137046	13	ABD32936-2
C	80	48.4	12.1	137046	13	ABD32936-2
C	81	48.4	12.1	210920	12	ADQ97123 Human can
C	82	48.2	12.1	1544	13	ADV1681 Rat card1
C	83	48.2	12.1	1544	14	ADV1681 Rat card1
C	84	48.2	12.1	118502	11	ACE11878 Rat surto
C	85	48.2	12.1	118502	11	ACE11878 Rat surto
C	86	48.2	12.1	197525	13	ABD33389 Human gen
C	87	48	12.0	5952	4	AAK61403 Human gen
C	88	48	12.0	32217	4	AAK61403 Human gen
C	89	48	12.0	91609	12	ADQ59473 Human can
C	90	47.8	12.0	378	2	AAV87853 EST clone
C	91	47.8	12.0	1329	2	AAV33199 Secreted

92	47.8	12.0	3035	5	ADG920677	Human ABC	XX	Cytostatic; Gene Therapy; breast cancer; RAD21; human; chromosome 8q24;
93	47.8	12.0	7378	5	ADG920614	Human ABC	KW	gene; single nucleotide polymorphism; SNP; ds.
94	47.8	12.0	7379	6	ABL58158	Human ABC	KW	
95	47.8	12.0	31477	11	ACN44096	Mouse gen	XX	
96	47.8	12.0	32359	11	ACN44720	Mouse gen	OS	Homo sapiens.
97	47.8	12.0	38564	13	ACN44732	Type II d	XX	
98	47.8	12.0	55008	4	AKR80495	Human imm	PH	Key
99	47.8	12.0	108316	10	ADG87336	Human GPC	FT	variation
100	47.8	12.0	164702	8	ACF62730	Cancer ba	FT	
101	47.8	12.0	164702	8	ADB20845	MRP1 Dase	FT	
102	47.8	12.0	164702	10	ADB87934	Human UGT	FT	
103	47.8	12.0	164702	10	ADB96917	Human MDR	FT	variation
104	47.8	12.0	164702	10	ADB92108	Human MDR	FT	
105	47.8	12.0	326014	6	ABK89296	Human gen	FT	variation
106	47.8	12.0	326014	12	ADQ94981	Human kin	FT	
107	47.6	11.9	426	14	ADM83621	MAP3f9 na	FT	
108	47.6	11.9	3657	10	ABZ22598	Human F13	FT	variation
109	47.6	11.9	47804	1	ACD86176	Human GPC	FT	
110	47.6	11.9	87311	11	ACN44792	Mouse gen	FT	
111	47.6	11.9	110000	12	ADG97266_0	Human can	FT	variation
112	47.6	11.9	220895	6	ABK84798	Human CDN	FT	
113	47.6	11.9	220895	13	ADBS2737	Drug cher	FT	
114	47.6	11.9	349881	10	ADC86642	Human GPC	FT	variation
115	47.6	11.9	349938	10	ADC87621	Human GPC	FT	
116	47.4	11.9	389	2	AAV90181	EST clone	FT	
117	47.4	11.9	4361	4	ABL04348	Drosophi1	FT	variation
118	47.4	11.9	13217	4	AAAS35088	DNA #38 e	FT	
119	47.4	11.9	13217	4	AAAS31477	Human DNA	FT	
120	47.4	11.9	13217	4	AAI63997	Human pol	FT	variation
121	47.4	11.9	13217	4	ABA06818	Human gen	FT	
122	47.4	11.9	13217	5	ABAI18587	Human ner	FT	
123	47.4	11.9	13217	5	AAAS29749	Human end	FT	variation
124	47.4	11.9	13217	6	ABQ66801	Human pol	FT	
125	47.4	11.9	13217	6	ABV84155	Human pol	FT	
126	47.4	11.9	13217	10	ADCI1088	Human DNA	FT	variation
127	47.4	11.9	13217	10	ADC46530	Human neo	FT	
128	47.4	11.9	13217	12	ADM24712	Human PRO	FT	
129	47.4	11.9	13220	4	AAAS35086	DNA #36 e	FT	variation
130	47.4	11.9	13220	4	AAAS31475	Human DNA	FT	
131	47.4	11.9	13220	4	AAI63995	Human pol	FT	
132	47.4	11.9	13220	4	ABA06816	Human gen	FT	variation
133	47.4	11.9	13220	5	ABAI18585	Human ner	FT	
134	47.4	11.9	13220	5	AAAS29747	Human end	FT	
135	47.4	11.9	13220	6	ABQ66799	Human pol	FT	variation
136	47.4	11.9	13220	6	ABV84153	Human pol	FT	
137	47.4	11.9	13220	10	ADCI1086	Human DNA	FT	
138	47.4	11.9	13220	10	ADC46528	Human neo	FT	variation
139	47.4	11.9	13220	12	ADM24710	Human PRO	FT	
140	47.4	11.9	73308	6	ABI66966	Lung canc	FT	
141	47.4	11.9	183999	4	AAI92831	Human ABC	FT	
142	47.4	11.9	183999	13	ADU76470	Human ABC	FT	variation
143	47.2	11.8	1068	6	ABBS6590	Mouse gen	FT	
144	47.2	11.8	6652	6	ABQ66988	Human ang	FT	
145	47.2	11.8	17967	6	ABL13015	Human imm	FT	variation
146	47.2	11.8	31477	11	ACN44096	Mouse gen	FT	
147	47.2	11.8	76341	12	ADM97422	Prostate	FT	
148	47.2	11.8	133632	11	ACN45054	Human gen	FT	variation
149	47.2	11.8	337022	12	ADQ59416	Human can	FT	
150	47	11.8	499	4	AAI86291	Human pol	FT	

ALIGNMENTS

RESULT 1

ID	ADP07906 standard; DNA; 83400 BP.
----	-----------------------------------

ADP07906;

26-AUG-2004 (First entry)

Human RAD21 gene, SEO ID 1,

DE

XX	Cytosaratic; Gene Therapy; breast cancer; RAD21; human; chromosome 8q24;
KM	gene; single nucleotide polymorphism; SNP; ds.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	variation
FT	location/Qualifiers
FT	replace(56,C)
FT	/*tag= a
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(7324,A)
FT	/*tag= b
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(7363,G)
FT	/*tag= c
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(9231,T)
FT	/*tag= d
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(10490,A)
FT	/*tag= e
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(11867,G)
FT	/*tag= f
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(13308,A)
FT	/*tag= g
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(13191,T)
FT	/*tag= h
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(13525,A)
FT	/*tag= i
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(13582,G)
FT	/*tag= j
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(15875,A)
FT	/*tag= k
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(17237,C)
FT	/*tag= l
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(18472,G)
FT	/*tag= m
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	18828
FT	/*tag= n
FT	/note= "The N at this position is either CTTTA or is deleted"
FT	replace(19170,C)
FT	/*tag= o
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(19349,T)
FT	/*tag= p
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(23869,A)
FT	/*tag= q
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(24235,A)
FT	/*tag= r
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(26691,A)
FT	/*tag= s
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(31373,T)
FT	/*tag= t
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(31979,G)
FT	/*tag= u
FT	/standard_name= "Single Nucleotide Polymorphism"
FT	replace(33472,C)
FT	/*tag= v

38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was determined by the technique of SBH (sequencing by hybridisation). Also included is a purified polypeptide comprising a sequence corresponding to a reading frame of the novel polynucleotide. The nucleic acid sequences are useful in diagnostics as expressed sequence tags (EST) for identifying expressed genes or for physical mapping of the human genome, in forensics, in assessing biodiversity, or in identifying mutations responsible for genetic disorders and other traits. The nucleotide sequences are also useful as hybridisation probes, as oligomers for PCR, for chromosome and gene mapping, in the recombinant production of protein, or in generating antisense DNA or RNA. The purified polypeptide is useful for generating antibodies specific for it. The present sequence is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?docid=20030073623

Sequence 463 BP, 117 A; 83 C; 87 G; 172 T; 0 U; 4 Other;

Query Match 14.0%; Score 56; DB 9; Length 463;

Best Local Similarity 62.3%; Pred. No. 3.2e-05; Mismatches 51; Indels 0; Gaps 0;

Matches 86; Conservative 1; Mismatches 51; Indels 0; Gaps 0;

106 GCAAGGCAAGTGAAGCCTACTTTTGTACTGTTGATATATATATATACATACACAC 165
343 GCACCCCACTTGAAAGAAATATAGTGAAGCTGTGATCTACACACACACACACAC 284
166 ACACACACACACACACATATTTTAACTGGGACATTAATATACACACCTATTCAG 225
283 ACACACACACACACACACAAATGAGTAAGACATTAATATGTAAGAAACCAAA 224
226 TAAAACTACTATCTCAG 243
223 AAAAATAAATGTAGAG 206

RESULT 6
ACN43862
ID ACN43862 standard; DNA; 141463 BP.

ACN43862;

18-NOV-2004 (first entry)

Human genomic sequence hCG21073.

Cytosbatic; carcinoma; lymphoma; cancer; human; gene; ss.

Homo sapiens.

MO2003073826-A2.

12-SEP-2003.

28-FEB-2003; 2003WO-US006235.

01-MAR-2002; 2002US-00087192.

(SAGR-) SAGRES DISCOVERY.

Morris DW;

WPI; 2003-328604/31.

Recombinant nucleic acid useful for diagnosis and treatment of carcinoma

comprises a nucleotide sequence.

Claim 1; SEQ ID NO 22; 0pp; English.

The present invention relates to novel DNA and protein sequences which are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcinoma Associated Protein (CAP); (iii) for screening of

a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a bioclip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA coding sequence. Note: This patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published

Sequence 141463 BP; 40336 A; 28306 C; 29237 G; 43584 T; 0 U; 0 Other;

Query Match 13.4%; Score 53.6; DB 11; Length 141463;

Best Local Similarity 56.2%; Pred. No. 0.00073; Mismatches 90; Indels 1; Gaps 1;

Matches 118; Conservative 1; Mismatches 90; Indels 1; Gaps 1;

133 TACTGTTGATATATATATATATATACATACACACACACACACACATATATTTAA 192
58906 TTCTTTAAATATATATATATATATATACATATACACACACACACATATATATA 58965
193 CTGGGACATTAATGTACACACCTATTCTAAAGTAAATCTACTATCTCAGAGTAAGA 252
58966 AAATATACACATAT-ATGTGTACAACTCTATATAAAATATATATGTTCACATGTA 59024
253 CATATCTCTGATATGTCATTAACTAGCAAGTAACTGCTTAACCTTATGCAAGTCACT 312
59025 TTTTAAACAGATGTTTTTTTAAAGCAATTAATTAATTTTATATTTTTCAGTTTCA 59084
313 TTCCCTACTATGAAACCTGAAGGTTGAA 342
59085 TTCCATATATATATATATATATGATATATATA 59114

RESULT 7
AAH41011/c
ID AAH41011 standard; DNA; 3455 BP.

AAH41011;

23-AUG-2001 (first entry)

Human metalloprotease related promoter SEQ ID 31.

Metalloprotease; human; aggrecanase; joint disease; osteoarthritis;

osteopathic; antiarthritic; promoter; ds.

Homo sapiens.

MO200134785-A1.

17-MAY-2001.

10-NOV-2000; 2000WO-JP007917.

11-NOV-1999; 99JP-00321740.

16-MAY-2000; 2000JP-00144020.

(YAMA) YAMANOUCHI PHARM CO LTD.

(KAZU-) KAZUSA DNA RES INST.

YamaJ N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;

WPI; 2001-343602/36;

Metalloprotease with aggrecanase activity for treating joint diseases

especially osteoarthritis.

Claim 11; Page 80-81; 85pp; Japanese.

This invention relates to a metalloprotease with aggrecanase activity. The invention includes protein and DNA sequences of the metalloprotease, vectors containing the DNA, host cells transformed by the vectors, and

The invention relates to novel foetal urogenital sinus-derived expressed sequence tags. The invention more specifically comprises an isolated polynucleotide comprising a nucleotide sequence containing a urogenital sinus-derived expressed sequence tag (EST) comprising g932, ug993, ug101, ug102, ug106, ug120, ug254, ug329, ug307, ug308, ug311, ug317, ug320, ug334, ug335, ug353, ug354, ug357, ug440, ug441, ug482, ug484, ug485, ug491, ug493, ug494, ug503, ug505, ug506, ug518, ug5186, ug5194. The urogenital sinus-derived expressed sequence tags have cytosstatic and anti-inflammatory activities. The genes containing urogenital sinus-derived ESTs encode oncofetal proteins that are useful for diagnosis of prostate disease and for monitoring the efficacy of a compound in clinical trials for the treatment of prostate disease. The ESTs can be used in assays to identify substances (e.g. triple helix forming oligonucleotides, antisense oligonucleotides, ribozymes or antibodies) useful for treating

The invention relates to cancer-associated proteins (CAP) and the cancer-associated (CA) nucleic acids encoding them. The invention also relates to a method for treating cancers involving administering to a patient an inhibitor of CAP, and a method of screening for anticancer activity in a potential drug involving providing a cell that expresses a CA gene, contacting a tissue sample derived from a cancer cell with an anticancer drug candidate and monitoring the effect of the anticancer drug candidate on expression of the CA gene. The CAP proteins are useful for detecting cancer associated with expression of a CAP protein in a test cell sample and for screening for a bioactive agent capable of modulating the activity of a CAP protein. The CA nucleic acids are useful for diagnosing cancer, involving determining the expression of a CA nucleic acid in a tissue. This sequence represents a murine CA gene of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pat_sequences](http://wipo.int/pub/published_pat_sequences)

Seq	Sequence	49914 BP;	12206 A;	12376 C;	12773 G;	12559 T;	0 U;	0 Other
Query Match		12.8%	Score 51.2;	DB 13;	Length 49914;			

PI Yamaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N,
XX
DR WPI; 2001-343602/36.
XX
PT Metalloprotease with aggrecanase activity for treating joint diseases
XX especially osteoarthritis.
XX
PS Claim 11; Page 76-77; 85pp; Japanese.

This invention relates to a metalloprotease with aggrecanase activity.
The invention includes protein and DNA sequences of the metalloprotease,
vectors containing the DNA, host cells transformed by the vectors, and
antibodies directed against the metalloprotease. The antibodies, protein
and DNA sequences can be used in the treatment and prevention of joint
diseases, particularly osteoarthritis. The treatment may result in
osteoprotic and antirheumatic activity. The present sequence represents a
promoter related to the metalloprotease of the invention

SQ Sequence 3467 BP; 819 A; 895 C; 893 G; 860 T; 0 U; 0 Other;

Query Match 12.8%; Score 51; DB 5; Length 3467;
Best Local Similarity 82.6%; Pred. No. 0.0013;
Matches 57; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

OY 121 CCTACCTTTTGTACTGTTGGATATATATATATATATACATACACACACACACACACA 180
DB 2418 CATTCATATATATATACATATCATATATATATATATATACATACACACACACACACACA 2359
OY 161 CAAATATT 189
DB 2358 CAATATAT 2350

22-JAN-2001 (first entry)

Human dual-specificity phosphatase-1 (DSP-1) gene.

DSP-1; dual-specificity phosphatase-1; human; cell proliferation; cell differentiation; cell survival; cell cycle; dephosphorylation; signal transduction; MAP-kinase; cancer; graft versus host disease; allergy; autoimmune disease; metabolic disease; therapy; chromosome 17; ds.

Homo sapiens.

Key Location/Qualifiers

exon 42..109

intron /tag= a

intron /number= 1

intron 110..20823

exon /tag= b

exon /number= 1

exon 20824..21034

exon /tag= d

exon /number= 2a

exon /note= "Alternative, extended version of exon 2"

exon /tag= c

exon 20824..20911

intron /number= 2

intron /tag= e

intron /number= 2

intron 21035..22327

intron /tag= f

exon /number= 2a

exon 22328..23309

exon /tag= g

exon /number= 3

CDS 22420..23016

W0200053636-A2.

14-SEP-2000.

08-MAR-2000; 2000MO-US006154.

08-MAR-1999; 99US-0123255P.

(CEPT-) CEPTYR INC.

Lucho RM, Wei B;

WPI; 2000-579365/54.

P-PSDB; AAB19602.

New isolated polypeptide having the sequence of dual-specificity phosphatase-1 (DSP-1) is useful for treating a patient with a disorder associated with DSP-1 activity e.g. cancer and autoimmune diseases.

Example 1; Fig 3A-J; 74pp; English.

The present sequence is that of the human dual-specificity phosphatase-1 (DSP-1) gene on chromosome 17. The gene was identified in genomic sequences obtained from an expressed sequence tag database screened with a conserved motif (see AAB19604) of known DSPs. DSP-1 dephosphorylates both phosphoserine/threonine and phosphotyrosine residues in DSP-1 substrates such as activated mitogen-activated protein kinase (MAP-kinase). DSP-1 has sequence homology to other MAP-kinase phosphatases. It is expressed at high levels in the human heart, testis and liver, and at lower levels in other tissues. Methods are provided for recombinant production of DSP-1 polypeptides, and for using DSP-1 polypeptides, antibodies and polynucleotides to detect DSP-1 expression, to screen for agents that modulate DSP-1 activity e.g. within a combinatorial library, and for using such agents to modulate cell proliferation, cell differentiation or cell survival, through modulation of pattern of gene

expression, apoptosis or cell cycle. In particular, the cell displays contract inhibition of cell growth, anchorage-dependent growth or an altered intercellular adhesion property, or is a cell present in a patient afflicted with a disorder associated with DSP-1 activity, such as cancer, graft-versus host disease, autoimmune disease, allergy, metabolic disease, abnormal cell growth, abnormal cell proliferation and abnormal cell cycle

Sequence 24000 BP; 6038 A; 5505 C; 5602 G; 6855 T; 0 U; 0 Other;

Query Match 12.8%; Score 51; DB 3; Length 24000;

Best Local Similarity 56.4%; Pred. No. 0.0023;

Matches 93; Conservative 1; Mismatches 71; Indels 0; Gaps 0;

94 GAAAGTAGGTATGACAGGCAAGTGAAGCTCTTTTGTACTGTTGATATATATAT 153

19220 GAGTGAAAGAAAGCCAGCAAAAAAACCACATATATATATATATATATATAT 19279

154 ACATACACACACACACACACACACACACACATATTTTAACGCGGACATTAATGACACC 213

19280 ATACACACACACACACACACACACACACACACACACACACCTTATATATATGTA 19339

214 ACCATTTCAAAGTAAACCTACTATCTCAAGAGTAAACACATTAAT 258

19340 GTATATGTGTACTATATATATATATATATATATATATATATATATATATAT 19384

RESULT 18

ID ABX95685/C

ABX95685 standard; DNA; 28854 BP.

ABX95685;

30-JUN-2003 (first entry)

Human gene encoding a protease of the metalloprotease family.

Human; ds; gene; chromosome 11; protease; metalloprotease; cancer; disintegrin; metalloprotease with thrombospondin motifs-1; cachexia; cytosolic; immunomodulator; SNP; single nucleotide polymorphism.

Homo sapiens.

Key Location/Qualifiers

variation replace(497,C)

variation /tag= q

variation /standard_name= "Single nucleotide polymorphism"

variation replace(649,T)

variation /tag= r

variation /standard_name= "Single nucleotide polymorphism"

variation replace(875,G)

variation /tag= s

CDS /standard_name= "Single nucleotide polymorphism"

2001..26854

product= "protease"

/transl_except= (pos:24100..24101,aa:Aaa)

2001..2957

/tag= b

exon /number= 1

2958..14516

/tag= c

intron /number= 1

variation replace(3332,A)

variation /tag= t

variation /standard_name= "Single nucleotide polymorphism"

variation replace(4208..4210,TG)

variation /tag= u

variation /standard_name= "Single nucleotide polymorphism"

variation replace(4773,A)

variation /tag= v

variation /standard_name= "Single nucleotide polymorphism"

variation replace(5009,G)

```

FT      /*tag= w
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(5597,G)
FT      /*tag= x
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(5664,C)
FT      /*tag= y
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(6323. .6324,TA)
FT      /*tag= z
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(6325,A)
FT      /*tag= aa
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(6564,G)
FT      /*tag= ab
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(10433. .10434,GGA)
FT      /*tag= ac
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(13184,G)
FT      /*tag= ad
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(13651,C)
FT      /*tag= ae
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(14517. .14517. .14649
FT      /*tag= d
FT      /number= 2
FT      /replace(14650. .15114
FT      /*tag= e
FT      /number= 2
FT      /replace(15114. .15282
FT      /*tag= f
FT      /number= 3
FT      /replace(15283. .15524
FT      /*tag= g
FT      /number= 3
FT      /replace(15471,G)
FT      /*tag= af
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(15525. .15808
FT      /*tag= h
FT      /number= 4
FT      /replace(15809. .22292
FT      /*tag= i
FT      /number= 4
FT      /replace(17255,T)
FT      /*tag= ag
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(17727,G)
FT      /*tag= ah
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(17833,G)
FT      /*tag= ai
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(18034. .18036,AA)
FT      /*tag= aj
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(18035,A)
FT      /*tag= ak
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(19687,A)
FT      /*tag= al
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(19371,G)
FT      /*tag= am
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(19478,C)
FT      /*tag= an
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(19803,A)
FT      /*tag= ao

```

```

FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(22293. .22470
FT      /*tag= j
FT      /number= 5
FT      /replace(22448,C)
FT      /*tag= ar
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(22448,A)
FT      /*tag= ag
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(22448,G)
FT      /*tag= ap
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(23952
FT      /*tag= k
FT      /number= 5
FT      /replace(23568,T)
FT      /*tag= as
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(23820,A)
FT      /*tag= at
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(23953. .24134
FT      /*tag= l
FT      /number= 6
FT      /replace(24135. .24240
FT      /*tag= m
FT      /number= 6
FT      /replace(24241. .24416
FT      /*tag= n
FT      /number= 7
FT      /replace(24417. .26079
FT      /*tag= o
FT      /number= 7
FT      /replace(25713. .25715,CC)
FT      /*tag= au
FT      /standard_name= "Single nucleotide polymorphism"
FT      /replace(26080. .26851
FT      /*tag= p
FT      /number= 8
FT      exon
FT      US2002086400-A1.
FT      04-JUL-2002.
FT      21-DEC-2000; 2000US-00741151.
FT      06-DEC-2000; 2000US-0251398P.
FT      (ZHUS/) ZHU S.
FT      (GUEG/) GUEGLER K.
FT      (WRS/) WEBSTER M.
FT      (DPR/) DI FRANCESCO V.
FT      (BEAS/) BEASLEY E M.
FT      Zhu S, Guegler K, Webster M, Di Francesco V, Beasley EM;
FT      WPI: 2003-401333/38.
FT      P-PSDB; ABU09520.
FT      Novel human protease protein related to metalloprotease subfamily and
FT      nucleic acid molecule encoding the protein for diagnosing, treating
FT      disease or condition mediated by the protease protein e.g. cancer,
FT      cachexia.
FT      Claim 4; Fig 3; 64pp; English.
FT      The invention relates to an isolated human protease peptide (a member of
FT      the disintegrin/metalloprotease with thrombospondin motifs-1 family), an
FT      allelic variant or orthologue encoded by a nucleic acid molecule that
FT      hybridises under stringent conditions to the opposite strand of the cDNA
FT      and gene appearing as (ABX95684 and ABX95685) or a fragment of having 10
FT      contiguous amino acids. Also included are the encoding nucleic acids
CC

```


cc (including their complements, allelic variants, orthologues or
cd fragments), an isolated antibody that selectively binds to the protease,
cd a gene chip comprising the nucleic acids, a transgenic non-human animal
cc comprising the nucleic acids, vectors and host cells producing the
cc protein. The protease is useful for identifying a modulator of the
cd expression of the protease or binds to it. The protease and the nucleic
cc acids are useful for treating, preventing and/or diagnosing disorders
cc such as cancer and cachexia. The protease and the nucleic acids are
cd further useful as a query sequence to perform a search against sequence
cd databases to identify other family members or related sequences. The

Query Match	12.8%	Score 51	DB 8	Length 28854
Best Local Similarity	82.6%	Pred. No. 0.0024		
Matches 57, Conservative			Mismatches 11	Indels 0
				Gaps 0

[illegible]

RESULT 19
ADA02927/c
ID ADA02927 standard; DNA; 92726 BP.

AC	ADA02927;
XX	
DT	06-NOV-2003 (first entry)

Mouse *Pik3r1* carcinoma associated gene, SEQ ID NO:1445

KM Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast
 KM prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
 KM gene; ds.

OS Mus sp.
XX
PN W02003057146-A2.

PD 17-JUL-2003.

26-DEC-2002; 2002WO-US041414.

PR 26-DEC-2001; 2001US-00035832.

PA (SAGR-) SAGRES DISCOVERY.

PI MORRIS DW;

WPI; 2003-587068/55.

PT New recombinant nucleic acid encoding carcinoma associated protein,
PT useful for preparing compositions for treating carcinomas.

PS Claim 1; SEQ ID NO 1445; 245pp; English.

The invention relates to recombinant carcinoma associated (CA) nucleic acid sequences from mouse and human (ADN01482-ADN03094), and to recombinant carcinoma associated proteins (CAP) encoded by them. The invention also encompasses expression vectors and host cells comprising a CA nucleic acid, a polypeptide (especially an antibody) that specifically binds to the protein, and a biolchip comprising CA nucleic acid or fragments thereof. The sequences of the invention were identified using oncogenic retroviruses, which insert into the genome of the host organism at random. Many of these do not carry transduced host oncogenes or pathogenic trans-acting viral genes, meaning that cancer incidence is a direct consequence of the effects of proviral integration into host protooncogenes. The CA nucleic acid sequences can be used to diagnose carcinoma (especially breast cancer, prostate cancer, lymphoma or leukemia) or a propensity to carcinoma by determination of the sequence

of a CA gene, or by determination of CA gene expression in particular tissues. CA nucleic acids, proteins and antibodies are also useful as therapeutic agents and in screening and evaluating drug candidates. The present sequence represents a specifically claimed murine CA nucleic acid sequence of the invention. Note: The complete sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences.

SQ Sequence 92726 BP; 23819 A; 19174 C; 20109 G; 27030 T; 0 U; 2594 Other;

Query Match	12.7%;	Score 50.8;	DB 9;	Length 92726;
Best Local Similarity	61.7%;	Pred. No. 0.0038;		
Matches 79;	Conservative 1;	Mismatches 48;	Indels 0;	Gaps 0

Oy	129	TTTGTACTGTTGGATTAATAATTATATACATACACACACACACACATATTT	188
Dd	77650	TATATATATATATATATATATATATATATACACACATACACACACACACACA	77599
Oy	189	TTAACTCGGCAGATAAATGTACACCACTTCGAAGTAAAACTATCTTCCAGAAGTA	248
Dd	77590	TTTACACATATGTGGCATATATAACAAAATTTGTGGATATATATACACACACACA	77533

```

Qy      249 AAGACATA 256
          | | | | |
Db      77530 CACACATA 77523

```

RESULT 20
ADB72665/c
ID ADB72665 standard; DNA; 92726 BP.

AC ADB72665 ;

DT 04-DEC-2003 (first entry)

DE Mouse Pik3r1 gene

KM mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KM cancer; neoplasm; adenocarcinoma; sarcoma; gene.

OS	Mus sp.
XX	
PN	W02003008583-A2.

PD 30-JAN-2003.

PF 26-DEC-2001; 2001WO-US051291.

PR 02-MAR-2001; 2001US-00798586.

PR 08-NOV-2001; 2001US-00052482.

PR 20-DEC-2001; 2001US-00034650.

PA (SAGR-) SAGRES DISCOVERY.

PI Morris DW, Engelhard EK;
 PI

(SAGR-) SAGRES DISCOVERY.

Morris DW, Engelhard EK;

WPI; 2003-239337/23.

New recombinant nucleic acid, useful for treating carcinomas, lymphomas, cancers, neoplasm, adenocarcinoma, or sarcomas.

Claim 1; SEQ ID NO 493; 2304pp; English.

CC The invention relates to a novel recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the 650 sequences fully defined
CC in the specification. A polynucleotide of the invention has cytostatic
CC activity, and may have a use in gene therapy, or in a vaccine. The
CC recombinant nucleic acids and polypeptides are useful for treating
CC carcinomas, e.g. lymphomas, cancers, neoplasms, adenocarcinoma, and
CC sarcomas. The present sequence represents a mouse gene of the invention.

PR	06-SEP-2000	2000US-0230437P.	PR	17-NOV-2000	2000US-0249218P.	PR	17-NOV-2000	2000US-0249218P.	PR	17-NOV-2000	2000US-0249218P.
PR	06-SEP-2000	2000US-0230438P.	PR	17-NOV-2000	2000US-0249219P.	PR	17-NOV-2000	2000US-0249219P.	PR	17-NOV-2000	2000US-0249219P.
PR	08-SEP-2000	2000US-0231242P.	PR	17-NOV-2000	2000US-0249245P.	PR	17-NOV-2000	2000US-0249245P.	PR	17-NOV-2000	2000US-0249245P.
PR	08-SEP-2000	2000US-0231243P.	PR	17-NOV-2000	2000US-0249264P.	PR	17-NOV-2000	2000US-0249264P.	PR	17-NOV-2000	2000US-0249264P.
PR	08-SEP-2000	2000US-0231244P.	PR	17-NOV-2000	2000US-0249265P.	PR	17-NOV-2000	2000US-0249265P.	PR	17-NOV-2000	2000US-0249265P.
PR	08-SEP-2000	2000US-0231413P.	PR	17-NOV-2000	2000US-0249297P.	PR	17-NOV-2000	2000US-0249297P.	PR	17-NOV-2000	2000US-0249297P.
PR	08-SEP-2000	2000US-0231414P.	PR	17-NOV-2000	2000US-0249299P.	PR	17-NOV-2000	2000US-0249299P.	PR	17-NOV-2000	2000US-0249299P.
PR	08-SEP-2000	2000US-0232080P.	PR	17-NOV-2000	2000US-0249300P.	PR	17-NOV-2000	2000US-0249300P.	PR	17-NOV-2000	2000US-0249300P.
PR	12-SEP-2000	2000US-0232081P.	PR	01-DEC-2000	2000US-0250160P.	PR	01-DEC-2000	2000US-0250160P.	PR	01-DEC-2000	2000US-0250160P.
PR	14-SEP-2000	2000US-0232397P.	PR	05-DEC-2000	2000US-0250391P.	PR	05-DEC-2000	2000US-0250391P.	PR	05-DEC-2000	2000US-0250391P.
PR	14-SEP-2000	2000US-0232398P.	PR	05-DEC-2000	2000US-0251030P.	PR	05-DEC-2000	2000US-0251030P.	PR	05-DEC-2000	2000US-0251030P.
PR	14-SEP-2000	2000US-0232399P.	PR	05-DEC-2000	2000US-0251988P.	PR	05-DEC-2000	2000US-0251988P.	PR	05-DEC-2000	2000US-0251988P.
PR	14-SEP-2000	2000US-0232400P.	PR	06-DEC-2000	2000US-0256719P.	PR	06-DEC-2000	2000US-0256719P.	PR	06-DEC-2000	2000US-0256719P.
PR	14-SEP-2000	2000US-0232401P.	PR	08-DEC-2000	2000US-0251479P.	PR	08-DEC-2000	2000US-0251479P.	PR	08-DEC-2000	2000US-0251479P.
PR	14-SEP-2000	2000US-0233063P.	PR	08-DEC-2000	2000US-0251865P.	PR	08-DEC-2000	2000US-0251865P.	PR	08-DEC-2000	2000US-0251865P.
PR	14-SEP-2000	2000US-0233064P.	PR	08-DEC-2000	2000US-0251866P.	PR	08-DEC-2000	2000US-0251866P.	PR	08-DEC-2000	2000US-0251866P.
PR	14-SEP-2000	2000US-0233065P.	PR	08-DEC-2000	2000US-0251869P.	PR	08-DEC-2000	2000US-0251869P.	PR	08-DEC-2000	2000US-0251869P.
PR	21-SEP-2000	2000US-0234223P.	PR	08-DEC-2000	2000US-0251989P.	PR	08-DEC-2000	2000US-0251989P.	PR	08-DEC-2000	2000US-0251989P.
PR	21-SEP-2000	2000US-0234274P.	PR	11-DEC-2000	2000US-02541907P.	PR	11-DEC-2000	2000US-02541907P.	PR	11-DEC-2000	2000US-02541907P.
PR	25-SEP-2000	2000US-0234997P.	PR	05-JAN-2001	2001US-0259678P.	PR	05-JAN-2001	2001US-0259678P.	PR	05-JAN-2001	2001US-0259678P.
PR	25-SEP-2000	2000US-0234998P.	XX	(HUMA-) HUMAN GENOME SCI INC.		XX	(HUMA-) HUMAN GENOME SCI INC.		XX	(HUMA-) HUMAN GENOME SCI INC.	
PR	26-SEP-2000	2000US-0235484P.	PA	Rosen CA, Barash SC, Ruben SW,		PA	Rosen CA, Barash SC, Ruben SW,		PA	Rosen CA, Barash SC, Ruben SW,	
PR	27-SEP-2000	2000US-0235834P.	PI	WPI, 2001-483426/52.		PI	WPI, 2001-483426/52.		PI	WPI, 2001-483426/52.	
PR	27-SEP-2000	2000US-0235835P.	XX	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,		XX	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,		XX	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PR	29-SEP-2000	2000US-0236377P.	DR	useful for preventing, diagnosing and/or treating cancers and metastasis.		DR	useful for preventing, diagnosing and/or treating cancers and metastasis.		DR	useful for preventing, diagnosing and/or treating cancers and metastasis.	
PR	29-SEP-2000	2000US-0236378P.	XX	Disclosure; SEQ ID NO 22626; 3071bp + Sequence Listing; English.		XX	Disclosure; SEQ ID NO 22626; 3071bp + Sequence Listing; English.		XX	Disclosure; SEQ ID NO 22626; 3071bp + Sequence Listing; English.	
PR	29-SEP-2000	2000US-0236386P.	XX	AAK54951 to AAK64702 encode the human immune/hematopoiet							

CC a mutation and is at least one of the mutations given in the
CC specification. The gene has a mutation in an intron. The gene has a
CC mutation in a promoter or regulatory region. The method is useful in
CC diagnosing a patient with Cornelia de Lange Syndrome. The current
CC sequence represents the genomic sequence of the NIPBL gene. This gene is
CC located on human chromosome 5 in the region designated p13.1-13.3.
XX
SQ Sequence 188056 BP; 58089 A; 31468 C; 34382 G; 64117 T; 0 U; 0 Other;
Query Match 12.5%; Score 49.8; DB 15; Length 188056;
Best Local Similarity 50.2%; Pred. No. 0.0086;
Matches 120; Conservative 1; Mismatches 118; Indels 0; Gaps 0;
QY 20 ATATTTCGAGCTCTTGGAATAGCTTAATAATCCTGCGCTACATGACATGCTGAT 79
DB 69597 ATTTAAACATTATTTTGAATTAACGAAATTTCAATATATGGAATATTTAGCTAATTA 69538
QY 80 ACTGTTTACCTCTGGAAGTAGATGATGAGGCAAGTGAAGCTTCTTTTGTACTGTT 139
DB 69537 CACCTTTCCTAATGTTAATTTTGTGCGTGGGGAAGGCTTAAACTAAGTACATAT 69478
QY 140 TGATATATATATATATACATACACACACACACACACACACATATTTTAACTCGGA 199
DB 69477 ATGTGTGTATATATATATATATACATACACACACACACATGATATATATACACAC 69418
QY 200 CATTAATGTACACACACCTATTCAAAGTAAACTACTCTCAGAAAGTAAAGACTAAT 258
DB 69417 ACAACACTGTTTGTGAGCAAGAAATTTTAATATATAGAAATTTAAGTAAAGAAAT 69359
RESULT 34
AD057023/C
ID AD057023 standard; DNA; 590 BP.
AC AD057023;
XX
XX 21-OCT-2004 (first entry)
DE Novel canine microarray-related DNA sequence SeqID8325.
XX
XX canine microarray; drug screening; toxicity assay;
KM environmental pollutant; cellular response; gene expression profile;
KM toxic response; liver necrosis; fatty liver disease;
KM protein adduct formation; hepatitis; dog; ds.
XX
XX Canis familiaris.
OS
XX WO2004063324-A2.
PN
XX 29-JUL-2004.
PB
XX
XX 05-MAY-2003; 2003WO-US013853.
PE
XX
XX 03-MAY-2002; 2002US-0377240P.
PR
XX
XX (GENE-) GENE LOGIC INC.
PA (PRTZ) PFIZER PROD INC.
XX
XX Digians JC, Porter M, Wei T;
PI
XX
XX WPI; 2004-561890/54.
DR
XX
XX New isolated nucleic acid molecule, useful for drug screening and
PT toxicity assays or for assessing the impact, including toxicity, of a
PT compound, pharmaceutical agent or environmental pollutant on a cell or
PT living organism.
XX
XX Claim 1; SEQ ID NO 8325; 41bp; English.
XX
XX This invention is related to a novel isolated canine nucleic acid
CC sequences and the construction of canine microarrays containing a
CC significant portion of the canine genome. The isolated canine nucleic
CC acid sequences of the invention may be useful for drug screening and

CC toxicity assays. The invention is therefore useful for assessing the
CC impact, including toxicity, of a compound, pharmaceutical agent or
CC environmental pollutant on a cell or living organism. The methods are
CC useful for detecting genes that are up- or down-regulated in canines in a
CC disease state. The sequences are useful as diagnostic agents or markers
CC to detect a cellular response in a sample individually or as part of a
CC gene expression profile. It is also useful as a target for agents that
CC modulate gene expression or activity. The database is useful for
CC producing electronic Northern blots that allow the user to determine the cell
CC type or tissue in which a given gene is expressed and to allow
CC determination of the abundance or expression level of a given gene in a
CC particular tissue or cell. The methods are useful for determining the
CC similarity of a toxic response to one or more individual compounds. The
CC methods are useful for predicting at least one toxic response or the
CC likelihood that a compound or test agent will induce various specific
CC pathologies such as those of the liver (liver necrosis, fatty liver
CC disease, protein adduct formation or hepatitis), those of the kidney,
CC heart, brain or testes, or other pathologies associated with at least one
CC of the toxins. The methods are also useful for predicting or elucidating
CC the potential cellular pathways influenced, induced or modulated by the
CC compound or test agent due to the similarity of the expression profile
CC compared to the profile induced by a known toxin. The present sequence is
CC that of a canine DNA sequence which was claimed for use during the
CC production of a canine microarray of the invention.
XX
SQ Sequence 590 BP; 201 A; 86 C; 84 G; 213 T; 0 U; 6 Other;
Query Match 12.4%; Score 49.4; DB 13; Length 590;
Best Local Similarity 56.0%; Pred. No. 0.0021;
Matches 89; Conservative 2; Mismatches 68; Indels 0; Gaps 0;
QY 68 GAGACTGTATCTGTTACTCTGGAAGTAGATGACAGGCAAGTGAAGCTTACT 127
DB 457 GATACCTGTTTATTTGATTCCTCTGGAATTAACAAAGCTTACTGATTTGTAC 398
QY 128 TTTTGTACTGTTGATATATATATATATATACATACACACACACACACATATAT 187
DB 397 ATATATATATGTTAT 338
QY 188 TTTTAACTCGGACATTAATATATATATATATATATATATATATATATATATAT 226
DB 337 GTCAAGAAAT 299
RESULT 35
ABK33962/C
ID ABK33962 standard; DNA; 4858 BP.
XX
XX ABK33962;
AC
XX
XX 18-JUN-2002 (first entry)
DT
XX
XX Human DNA for staging of Astrocytomas, complement, #23.
DE
XX
XX Human; de; astrocytoma; cytostatic; staging; cysteine methylation; CpG;
KM bisulphite; brain tissue; MALDI; ESI; electron spray mass spectrometry;
KM matrix assisted laser desorption/ionization mass spectrometry.
XX
XX Homo sapiens.
OS
XX
XX WO200202808-A2.
PN
XX
XX 10-JAN-2002.
PD
XX
XX 02-JUL-2001; 2001WO-EP007538.
PE
XX
XX 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
XX (BFIG-) EPIGENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K;

WPI; 2002-171649/22.

Novel chemically modified genomic DNA sequences, useful in the characterization, classification, differentiation, grading, staging, treatment and/or diagnosis of astrocytomas or predisposition to astrocytomas.

Claim 1; SEQ ID NO 46; 37bp; English.

The invention relates to a nucleic acid comprising a sequence (I) of at least 18 bases in length of a segment of chemically pre-treated genomic DNA which has any one of the sequences of (ABX30319-ABX34032) or its complement. Also included are an oligonucleotide or peptide nucleic acid (or set thereof) of at least 9 nucleotides which hybridises to (I), primers for (I), probes for detecting cytosine methylation or single-nucleotide polymorphisms (SNP) in (I), an array of oligomers or peptide nucleic acids for analysing diseases associated with the methylation states of the CpG dinucleotides of (I). The array is useful for determining genetic and/or epigenetic parameters, classification, differentiation, grading, staging, treatment and/or diagnosis of astrocytomas, or the predisposition to astrocytomas by analysing cytosine methylations, involves obtaining a biological sample containing genomic DNA, extracting the genomic DNA, converting cytosine bases which are unmethylated at the 5-position, in the genomic DNA sample, to uracil or another base which is dissimilar to cytosine in terms of hybridisation behaviour, by chemical treatment and amplifying chemically pre-treated genomic DNA fragments using the array and a polymerase, where the amplification carries a detectable label. The method further involves identifying methylation status of one or more cytosine positions, and analysing methylation status of the cytosine positions by reference to one or more data sets. The genomic DNA is chemically treated by using a biolipite, hydrogen sulphite or disulphite. The amplification step amplifies DNA which is of particular interest in astrocytoma or brain tissue, based on the specific genomic methylation status of brain tissues, as opposed to background DNA. The amplification carries a fluorescent label or radionuclide. Optionally, the labels of the amplicons are detachable molecule fragments having a typical mass which are detected in a mass spectrometer. The fragments of chemically pre-treated genomic DNA to be amplified, have a single positive or negative charge for a better detectability in the mass spectrometer. Preferably, the amplicons or fragments of the amplicons are detected by matrix assisted laser desorption/ionization mass spectrometry (MALDI) or using electron spray mass spectrometry (ESI). The present invention is one of the chemically pre-treated reference DNA samples of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at <http://publ.int/publ/sequences>

Sequence 4858 BP; 995 A; 210 C; 1224 G; 2429 T; 0 U; 0 Other;

Query Match 12.4%; Score 49.4; DB 6; Length 4858;
Best Local Similarity 52.8%; Pred. No. 0.0039;
Matches 104; Conservative 1; Mismatches 92; Indels 0; Gaps 0;

```

QY 90 TCTGGAAGTAGTAGGAGGCAAGTGAAGGCTACCTTTTGTGACTGTTGATATATAT 149
DB 3711 TTTTAAACATCCAAACAAACAAAAACACATCTCTACACACACACACACACAC 3652
QY 150 ATATACATACACACACACACACACACACACACACATATTTTAACTCGGACATTAATGTA 209
DB 3651 ACACACACACACACACACACACACACACACACACACAAATTAAGTATTAATACATA 3592
QY 210 CACCACTATTCAAGTAAAGTAACTATCTCAGAGTAAAGACATATCTGCTATTGT 269
DB 3591 AACCTATATATCCCACTATTAAACAAACAAAACTAAAAAATTAATA 3532
QY 270 CATTAAGTACGAGTTAA 286
DB 3531 CCTTAACAAAAAATTA 3515

```

RESULT 36
ADA20413/c

ADA20413 standard; DNA; 4858 BP.

ADA20413;

20-NOV-2003 (first entry)

Prostate tumour related genomic DNA complement sample #39.

Cytosine; gene therapy; genetic marker; epigenetic parameter;

classification; differentiation; diagnosis; prostate tumour;

prostate cancer; cytosine methylation; uracil;

single nucleotide polymorphism; SNP; prostate carcinoma; ss.

Homo sapiens.

W02002103042-A2.

27-DEC-2002.

14-JUN-2002; 2002MO-EP006605.

14-JUN-2001; 2001DE-01028508.

(EPIG-) EPIGENOMICS AG.

Dietler J, Model F, Adorjan P;

WPI; 2003-167536/16.

Determining genetic and/or epigenetic parameters, useful for the PT classification, differentiation and/or diagnosis of prostate tumors or a PT predisposition to prostate cancer, comprises analyzing cytosine methylation.

Claim 28; Page 302-304; 376bp; English.

The invention relates to a method of determining genetic and/or epigenetic parameters for the classification, differentiation and/or diagnosis of prostate tumors or the predisposition to prostate cancer, by analysing cytosine methylation in a sample of genomic DNA. The method comprises chemically treating unmethylated cytosine bases at the 5-position to uracil or another base, which is dissimilar to cytosine in terms of hybridization behaviour, followed by amplifying at least one fragment of the chemically pre-treated genomic DNA using sets of primer oligonucleotides and a polymerase. The oligomers or probes derived from CC them are useful for detecting the methylation state of all CpG CC dinucleotides and/or single nucleotide polymorphisms (SNPs) in a CC chemically pre-treated genomic DNA. They are all useful for treating CC prostate carcinoma. This sequence represents a fragment of genomic DNA CC used in the method of the invention.

Sequence 4858 BP; 995 A; 210 C; 1224 G; 2429 T; 0 U; 0 Other;

Query Match 12.4%; Score 49.4; DB 8; Length 4858;
Best Local Similarity 52.8%; Pred. No. 0.0039;
Matches 104; Conservative 1; Mismatches 92; Indels 0; Gaps 0;

```

QY 90 TCTGGAAGTAGTAGGAGGCAAGTGAAGGCTACCTTTTGTGACTGTTGATATATAT 149
DB 3711 TTTTAAACATCCAAACAAACAAAAACACATCTCTACACACACACACACACAC 3652
QY 150 ATATACATACACACACACACACACACACACACACATATTTTAACTCGGACATTAATGTA 209
DB 3651 ACACACACACACACACACACACACACACACACACACAAATTAAGTATTAATACATA 3592
QY 210 CACCACTATTCAAGTAAAGTAACTATCTCAGAGTAAAGACATATCTGCTATTGT 269
DB 3591 AACCTATATATCCCACTATTAAACAAACAAAAATTAATA 3532
QY 270 CATTAAGTACGAGTTAA 286
DB 3531 CCTTAACAAAAAATTA 3515

```


[illegible]

CC	The cdk modulator is preferably N-5-(1,1-Dimethyl-2-
CC	oxazolylo[1methy]l-2-thiazoyl-4-piperidine carboxamide, 0.5-L-
CC	tartaric acid salt. Note: The sequence data for this patent did not form
CC	part of the printed specification, but was obtained in electronic format
CC	directly from WIPO at ftp.wipo.int/pub/published_pat_sequences. This
CC	sequence encodes a biomarker used in the method of the invention.
XX	
SQ	Sequence 154394 BP; 43162 A; 34235 C; 34150 G; 42847 T; 0 U; 0 Other;
Query Match	12.4%; Score 49.4; DB 14; Length 154394;
Best Local Similarity	72.9%; Pred. No. 0.01;
Matches	62; Conservative 1; Mismatches 22; Indels 0; Gaps 0;
OY	127 TTTTGTACTGTTGATATATATATATACCTAACACACACACACACACACAYATA 186 12277 TATGTGTGTATATATATATATATACACACACACACACACACACACACAGA 12218
Dd	187 TTTTAACTCGGGCATTAATGTCA 211 12217 TTCTGGGCTAAGATATATTATCA 12193
Db	
RESULT 41	
ID	ADZ80338/c
ADZ80338	standard; DNA; 154394 BP.
AC	ADZ80338;
XX	
XX	14-UU-2005 (first entry)
DT	
XX	
DE	Human nuclear transcription factor Y alpha, gene.
KW	Nuclear transcription factor Y; ds; gene; genetic marker;
KM	gene expression; cancer; tumor; neoplasm; cytostatic; diarrhea;
KW	antidiarrhetic; gastrointestinal diseases; epoehione B.
XX	
OS	Homo sapiens.
PN	WO2005039573-A2.
PD	06-MAY-2005.
PF	05-OCT-2004; 2004WO-EPO11122.
PR	06-OCT-2003; 2003US-0508973P.
PA	(NOVS) NOVARTIS AG.
PA	(NOVS) NOVARTIS PHARMA GMBH.
Pf	
PI	Wolfgang CD;
DR	WPI: 2005-355754/36.
PT	Predicting diarrhoea in subject to whom microtubule stabilizing agent is
PT	to be administered, based on gene expression profile of subject, that
PT	predicts occurrence of diarrhoea following administration of agent.
PS	Example: SEQ ID NO 17; 124pp; English.
XX	
XX	The invention relates to predicting (M) diarrhea in a subject to whom a
CC	microtubule stabilising agent is to be administered, involves (a)
CC	obtaining the gene expression profile of the subject, where the gene
CC	expression profile comprises the gene expression pattern of one or more
CC	genes that are predictive of the occurrence of diarrhea in subject
CC	following administration of microtubule stabilizing agent, determining
CC	whether the subject is at risk for diarrhea from the administration of
CC	the microtubule stabilizing agent, (b) determining whether the subject
CC	expresses the Diego blood type and determining whether the subject is at
CC	risk for diarrhea following the administration of microtubule stabilizing
CC	agent, or (c) determining whether the subject has a lower than normal
CC	hematological levels as determined by hematological assays chosen from
CC	hematocrit and hemoglobin levels, and determining whether the subject is
CC	at risk for diarrhea following the administration of the microtubule

Cy	Cystostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.
Ox	
Kv	Homo sapiens.
Xx	
Pn	WO2004060304-A2.
Xx	
Pd	22-Jul-2004.
Xx	
Pf	22-Dec-2003; 2003WO-US041389.
Xx	
Pr	27-Dec-2002; 2002US-00330773.
Xx	
Pa	(SAGR-) SAGRES DISCOVERY INC.
Xx	
Pt	Morris DW, Malandro MS;
Dk	WPI, 2004-543781/52.
Xx	
Pt	New isolated cancer associated nucleic acids comprising at least 10
Lt	contiguous nucleotides, useful for diagnosing, preventing and/or treating
Xx	cancers such as leukemia and lymphoma.
Xx	
Pg	Claim 1; SEQ ID NO 159; 19pp; English.
Xx	
Cc	The present invention relates to cancer associated sequences (ADG97025-
Cc	ADD98004). The sequences are useful for the diagnosis, prevention and/or
Cc	treatment of cancer, such as leukemia and lymphoma. Note: The sequence
Cc	data for this patent did not form part of the printed specification, but
Cc	was obtained in electronic format directly from WIPO at
Cc	ftp.wipo.int/pub/published_pat_sequences.
Sq	Sequence 138627 BP; 33083 A; 34745 C; 35023 G; 35281 T; 0 U; 495 Other;
Xx	
Qy	Query Match 12.3%; Score 49; DB 12; Length 138627;
Dk	Best Local Similarity 82.1%; Pred. No. 0.013; Mismatches 11; Indels 0; Gaps 0;
Xx	
Oy	142 ATATATATATATACATCACACACACACACACACACATTTTACTCGGGACA 201 53280 ATATATATATACACACACACACACACACACACACACATATATATATTAGACA 53221
Oy	202 TAAATGT 208 53220 TAAATTT 53214
Dk	
Rl	RESULT 46 ABK28343/c ID ABK28343 standard; DNA; 10138 BP. Xx AC Xx AG Xx ABK28343; Xx Xx 23-APR-2002 (first entry) Dk DE DNA transcription associated genomic DNA #109. Xx Xx Xx DNA transcription associated gene; peptide nucleic acid; RNA-oligomer; Km PNA; cytosine methylation state; SNP; retroviral infection; gene; ds; Km single nucleotide polymorphism; adenosine deaminase deficiency; cancer; Km vital infection; Sezary syndrome; haematological disorder; tuberculosis; Km immunologic disorder; Werner syndrome; developmental disorder; Km psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis; Km neurodegenerative disorder; Waardenburg syndrome; Niemann-Pick disease; Km myelodysplastic syndrome; myocardial infarction; hypertension; arthritis; Km angiodysplasias; congenital heart disease; HDR syndrome; gene therapy; Km polyglutamine disorder; solid tumour. Xx Xx Unidentified. Ox Xx WO200192565-A2. Xx EN Pd 06-DEC-2001.

[illegible]

ID		ACN43848 standard; DNA; 40633 BP.
XX	AC	ACN43948;
XX	DY	18-NOV-2004 (first entry)
XX	DE	Mouse genomic sequence MCG10049.
XX	KM	Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX	Mus musculus.	
PX	MO2003073826-A2.	
PN	12-SEP-2003.	
PD	28-FEB-2003; 2003WO-US006235.	
PF	01-MAR-2002; 2002US-00087192.	
PR	(SAGR-) SAGRES DISCOVERY.	
PA	Morris DW;	
PI	WPI; 2003-328604/31.	
DR	Recombinant nucleic acid useful for diagnosis and treatment of carcinoma comprises a nucleotide sequence.	
XZ	Claim 1; SEQ ID NO 1; Opp; English.	
XX	The present invention relates to novel DNA and protein sequences which are associated with carcinomas. The sequences are useful for:	(i) screening drug candidates;
CC	(ii) for screening of bioactive agent capable of binding to Carcinoma Associated Protein (CAP);	(iii) for screening of a bioactive agent capable of modulating the activity of CAP;
CC	(iv) for evaluating the effect of a candidate carcinoma drug;	(v) for diagnosing carcinoma;
CC	(vi) for inhibiting the activity of CAP;	(vii) for treating carcinoma;
CC	(viii) for neutralizing the effect of CAP;	(ix) as a biophop;
CC	(x) for diagnosing carcinoma or a propensity to carcinoma; and	(xi) for determining Carcinoma Associated (CA) gene copy number.
CC	In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including Lymphoma. The present sequence is one such CA coding sequence. Note: This patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published	
SQ	Sequence 40633 BP; 10555 A; 8795 C; 8589 G; 11134 T; 0 U; 1560 Other;	
Oy	Query Match	12.2%; Score 48.8; DB 11; Length 40633;
Db	Best Local Similarity	29.9%; Pred. No. 0.01;
Oy	Matches 59; Conservative 1; Mismatches 137; Indels 0; Gaps 0;	
Oy	17 TATAATTTTACGAGCTTCGTGAATAAGCTTAATAACTCCTGGCTGCACATGAGAAGTGST 76 	
DB	TATATTATATTANN 31126 	
Oy	77 GATACTGGTACCCTCTGGAAAGTAGATGACGGACAAGTGAAGGCCCTACTTTTTGTA CT 136 NN 31186	
Oy	137 GTTGATATATATATATATACATPACACACACACACACACACACAYATATTTACTCG 196 	
Db	31187 NNNNNNNNNNTATATATATATACACACACACACACACATATATATGAAGTAAGTA 31246 	
Oy	197 GGACATAAATGTACACC 213 	
Db	31247 TTACTTTAATGATTAAC 31263 	
<hr/>		
RESULT 48		
ABD32966_01/C		
Continuation (2 of 11) of ABDJ32966 from base 100001 (mouse cancer-associated genomic DN		
MP Sequence split into 11 fragments LOCUS ABDJ32966 Accession Abdj32966		

[illegible]

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 17, 2006, 16:57:45 ; Search time 35973 Seconds

(without alignments)
621.792 Million cell updates/sec

Title: SEQ1-33670G
Perfect score: 399.2
Sequence: 1 aagaatgctcttctctata.....taactcagatcagaagcgag 400

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues
Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 150 summaries

Database :

EST:*
1: gb_seq1:*
2: gb_seq3:*
3: gb_seq4:*
4: gb_seq5:*
5: gb_seq6:*
6: gb_seq7:*
7: gb_seq8:*
8: gb_seq9:*
9: gb_seq10:*
10: gb_seq11:*
11: gb_seq12:*
12: gb_seq13:*
13: gb_seq14:*
14: gb_seq15:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	217.6	54.5	681	4 BX486675	DKX2P6860
2	57.6	14.4	882	14 CT350937	Sus scrofa
3	54.2	13.6	1204	14 AG390615	Mus muscu
4	53.8	13.5	655	4 CB940690	IPCCJx14
5	53.8	13.5	868	14 CNO3018554	Tetradodon
6	53.6	13.4	612	14 FRO018554	Alou11449 F. rubripe
7	53.6	13.4	779	9 DR106610	DR106610 JH1118H03
8	53.4	13.4	491	11 AQ407117	HS_5076_A
9	53.4	13.4	757	11 BZ217953	BZ217953 CH230-518
10	53.4	13.3	650	10 DM252261	UI-S-GBO-
11	52.6	13.2	602	14 CT063242	Sus scrofa
12	52.6	13.2	768	14 CT376921	Sus scrofa
13	52.6	13.2	796	11 BH116000	RPCI-24-2
14	52.6	13.2	855	14 CT035778	Sus scrofa
15	52.6	13.2	1015	12 CL121100	ISB1-79P1
16	52.4	13.1	315	8 CN674247	A0946A09-
17	52.4	13.1	521	8 CN676171	A0976B03-
18	52.4	13.1	541	11 AQ353718	CITB1-E1-
19	52.4	13.1	619	11 AQ779979	HS_3113_A

20	52.4	13.1	736	14 AG561036	AG561036 Mus muscu
21	52.4	13.1	2558	6 AK142564	AK142564 Mus muscu
22	52.2	13.1	494	11 AQ209289	AQ209289 HS_3237_B
23	52.2	13.1	660	3 BQ188992	UI-E-E1-
24	52.2	13.1	665	3 BQ184399	BQ184399 UI-E-E1-
25	52.2	13.1	820	10 DM264376	DM264376 UI-S-GN1-
26	52.2	13.1	950	13 DU046739	DU046739 159926 To
27	52	13.0	448	11 AQ068109	AQ068109 nbdb0021B
28	52	13.0	799	14 CR023958	CR023958 Forward B
29	52	13.0	887	5 CD360167	CD360167 AGENCOURT
30	51.8	13.0	345	13 CL440749	CL440749 PSTV001.H
31	51.8	13.0	501	2 B1503732	B1503732 B170022A
32	51.8	13.0	649	13 CM904996	CM904996 RPCI42.14
33	51.6	12.9	528	4 CB940280	IPCCJx14
34	51.6	12.9	694	14 BX162787	BX162787 Dario rer
35	51.6	12.9	827	14 BX128732	BX128732 Dario rer
36	51.4	12.9	528	11 BH343871	BH343871 CH230-41N
37	51.4	12.9	777	11 BZ250267	BZ250267 CH230-353
38	51.2	12.8	178	11 AZ514487	AZ514487 1M0361P24
39	51.2	12.8	541	11 AZ703674	AZ703674 RPCI-23-2
40	51.2	12.8	750	14 CT454907	CT454907 Sus scrofa
41	51	12.8	607	10 DM340951	DM340951 PP_LEC000
42	51	12.8	607	11 AQ326274	AQ326274 nbdb0028M
43	51	12.8	652	9 DN877029	DN877029 nae10901.
44	51	12.8	672	11 AQ923572	AQ923572 RPCI-23-2
45	51	12.8	835	14 CT398920	CT398920 Sus scrofa
46	51	12.8	1072	10 DM594541	DM594541 CGK113-F0
47	51	12.8	1381	9 DN738453	DN738453 CNB98-F11
48	51	12.7	1722	6 AK139565	AK139565 Mus muscu
49	50.8	12.7	469	14 DE117100	DE117100 Oryzias 1
50	50.8	12.7	493	11 AZ073117	AZ073117 RPCI-23-4
51	50.8	12.7	599	11 AQ262841	AQ262841 CITB1-E1-
52	50.8	12.7	714	14 BX176458	BX176458 Dario rer
53	50.8	12.7	732	13 DU341743	DU341743 109831304
54	50.8	12.7	770	14 CR830901	CR830901 GROMAA63B
55	50.8	12.7	879	13 CZ992356	CZ992356 209924 To
56	50.8	12.7	1175	4 BX363884	BX363884 BX363884
57	50.6	12.7	488	11 AQ597715	AQ597715 HS_2090_B
58	50.6	12.7	590	11 AZ012541	AZ012541 RPCI-23-3
59	50.6	12.7	693	14 AG171472	AG171472 Pan trogl
60	50.6	12.7	825	14 CT053907	CT053907 Sus scrofa
61	50.6	12.7	892	14 CT354182	CT354182 Sus scrofa
62	50.4	12.6	274	3 BP510846	BP510846 BP510846
63	50.4	12.6	525	11 B64703	B64703 CIT-HSP-202
64	50.4	12.6	606	14 DR36J11T	DR36J11T Dario rer
65	50.4	12.6	607	14 CR880635	CR880635 Sus scrofa
66	50.4	12.6	623	14 BX980437	BX980437 Forward s
67	50.4	12.6	711	11 AZ337912	AZ337912 1M0068G18
68	50.4	12.6	780	14 CT048379	CT048379 Sus scrofa
69	50.4	12.6	2533	6 AK032764	AK032764 Mus muscu
70	50.4	12.6	5305	6 AK143960	AK143960 Mus muscu
71	50.2	12.6	540	13 CM353832	CM353832 fdbb001f0
72	50.2	12.6	608	10 DT448882	DT448882 JGI_CABK8
73	50.2	12.6	633	10 DT448883	DT448883 JGI_CABK8
74	50.2	12.6	674	14 AG034856	AG034856 Pan trogl
75	50.2	12.6	835	9 CX823483	CX823483 JGI_CAK3
76	50.2	12.6	1001	3 BQ941666	BQ941666 AGENCOURT
77	50	12.5	274	11 AZ517401	AZ517401 RPCI-11-2
78	50	12.5	617	12 CG976896	CG976896 CH240.167
79	50	12.5	619	11 AQ342672	AQ342672 JGI_CABK8
80	50	12.5	732	13 CM566830	CM566830 OA_ABA011
81	50	12.5	766	14 AG496779	AG496779 Mus muscu
82	50	12.5	814	12 CC905634	CC905634 t024r22ba
83	50	12.5	814	14 CT319339	CT319339 Sus scrofa
84	50	12.5	933	10 DM662837	DM662837 CNB316-GO
85	50	12.5	952	13 CZ941129	CZ941129 323146 To
86	49.8	12.5	563	12 BZ599380	BZ599380 WRADP43TF
87	49.8	12.5	634	14 BX157385	BX157385 Dario rer
88	49.8	12.5	646	14 AG065044	AG065044 Pan trogl
89	49.8	12.5	750	3 BQ002660	BQ002660 UI-H-E11-
90	49.8	12.5	751	4 AG603288	AG603288 Mus muscu
91	49.8	12.5	763	13 CL987558	CL987558 ZMWBE000
92	49.8	12.5	787	10 DT292845	DT292845 JGI_CAW4

C	93	49.8	12.5	803	13	CZ979367	CZ979367 197206 To
C	94	49.6	12.4	468	1	AL709462	AL709462 DKFZp686D
C	95	49.6	12.4	500	7	BE296556	BE296556 601174057
C	96	49.6	12.4	527	1	AJ424934	AJ424934 AJ424934
C	97	49.6	12.4	550	14	CR223106	CR223106 Forward s
C	98	49.6	12.4	586	14	CR260017	CR260017 Forward s
C	99	49.6	12.4	588	11	AQ721188	AQ721188 HS 5564 A
C	100	49.6	12.4	626	11	BZ214131	BZ214131 CH230-426
C	101	49.6	12.4	638	14	CR117368	CR117368 Forward s
C	102	49.6	12.4	662	14	DUE55930	DUE55930 OG AB8005
C	103	49.6	12.4	702	14	CR158437	CR158437 Forward s
C	104	49.6	12.4	758	14	AG522628	AG522628 Mus muscu
C	105	49.6	12.4	811	14	DUE60221	DUE60221 66448 Tom
C	106	49.6	12.4	829	14	CT406476	CT406476 Sus scrofa
C	107	49.6	12.4	906	14	CR235414	CR235414 Forward s
C	108	49.4	12.4	239	14	DE236382	DE236382 Tifolium
C	109	49.4	12.4	258	14	BX222469	BX222469 Danio rer
C	110	49.4	12.4	267	11	AZ121960	AZ121960 RPI-23-1
C	111	49.4	12.4	290	3	BQ989052	BQ989052 OGF16J24.
C	112	49.4	12.4	292	14	CT427657	CT427657 Sus scrofa
C	113	49.4	12.4	377	12	BZ930241	BZ930241 CH240.35J
C	114	49.4	12.4	481	7	BF703922	BF703922 M1-P-A5-a
C	115	49.4	12.4	532	9	DN871044	DN871044 nad1.0h06.
C	116	49.4	12.4	561	10	DV722099	DV722099 RVL13435
C	117	49.4	12.4	570	11	BH084190	BH084190 RPI-24-3
C	118	49.4	12.4	572	12	CE736719	CE736719 L19T-988-
C	119	49.4	12.4	626	14	BX687247	BX687247 BX687247
C	120	49.4	12.4	632	14	BX128212	BX128212 Danio rer
C	121	49.4	12.4	656	12	CE784742	CE784742 L19T-988-
C	122	49.4	12.4	662	9	DN744924	DN744924 GL-CI-171
C	123	49.4	12.4	763	14	CT1210091	CT1210091 Sus scrofa
C	124	49.4	12.4	781	14	CT189540	CT189540 Sus scrofa
C	125	49.4	12.4	789	11	BZ274753	BZ274753 Sus scrofa
C	126	49.4	12.4	860	11	CR274753	CR274753 Sus scrofa
C	127	49.4	12.4	947	14	CNS02MB6	CNS02MB6
C	128	49.4	12.4	981	12	CL054028	CL054028 Tetradon
C	129	49.2	12.3	223	12	CE396292	CE396292 L19T-988-
C	130	49.2	12.3	269	7	BE501814	BE501814 hm35e11.x
C	131	49.2	12.3	377	10	DM451534	DM451534 HRA80042
C	132	49.2	12.3	435	11	AQ041986	AQ041986 CTT-HSP-2
C	133	49.2	12.3	531	12	CL193434	CL193434 104_417-1
C	134	49.2	12.3	640	11	AZ952702	AZ952702 2M021E11
C	135	49.2	12.3	658	11	AZ643106	AZ643106 1M0506L05
C	136	49.2	12.3	702	14	CR048186	CR048186 Forward s
C	137	49.2	12.3	746	14	AG606779	AG606779 Mus muscu
C	138	49.2	12.3	753	11	AQ895904	AQ895904 HS_3176-B
C	139	49.2	12.3	777	11	AQ742337	AQ742337 HS_5383-B
C	140	49.2	12.3	780	14	CR267037	CR267037 Reverse s
C	141	49.2	12.3	783	14	BX971441	BX971441 Reverse s
C	142	49.2	12.3	785	14	AG349596	AG349596 Mus muscu
C	143	49.2	12.3	808	14	CR007296	CR007296 Reverse s
C	144	49	12.3	349	4	BX955892	BX955892 DKFZp7811
C	145	49	12.3	564	11	AZ329878	AZ329878 1M0054L08
C	146	49	12.3	564	11	AZ515186	AZ515186 1M0054L08
C	147	49	12.3	589	13	DUI175448	DUI175448 109855324
C	148	49	12.3	653	11	AZ603140	AZ603140 1M0422G17
C	149	49	12.3	725	14	AG497075	AG497075 Mus muscu
C	150	49	12.3	762	14	AG399545	AG399545 Mus muscu

ALIGNMENTS

RESULT 1
LOCUS BX486675 681 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp686O06252.1 686 (synonym: hlc3) Homo sapiens cDNA clone
ACCESSION BX486675
VERSION BX486675.1 GI:31950573
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Bahr, A., Lauber, J., Mewes, H.W., Weill, B., Amid, C., Oanger, A., Fodor, G., Han, M., and Wiemann, S.
TITLE EST (Bahr, A., Lauber, J., Mewes, H.W., Weill, B., et al.)
JOURNAL Unpublished (2003)
COMMENT
CONTACT: MIPS
MIPS
Ingolstaedter Landstr. 1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
This clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project.
No 5' sequence available.
This clone (DKFZp686O06252) is available at the RZPD in Berlin.
Please contact the RZPD: Resourcentrum, Heuberger 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1..681
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686O06252"
/dev_stage="adult"
/lab_host="DH10B"
/note="11b="686 (synonym: hlc3)"
/note="Vector: pTriblex2; Site_1: SfilA; Site_2: SfilB; cDNA-collection"

ORIGIN

Query Match 54.5%; Score 217.6; DB 4; Length 681;
Best Local Similarity 96.8%; Pred. No. 2.6e-44;
Matches 242; Conservative 2; Mismatches 1; Indels 5; Gaps 2;
QY 1 AAGATGCTCTTCTATATATTTACGAGTCTTGATAGCTTAATAATACCTGGCG 60
DB 249 AAGATGCTCTTCTATATATTTACGAGTCTTGATAGCTTAATAATACCTGGCG 190
QY 61 TACACATGACGTGTGATACCTGTTACCTGGAAGATGATGACGAGGCAAGTGAAG 120
DB 189 TACACATGACGTGTGATACCTGTTACCTGGAAGATGATGACGAGGCAAGTGAAG 130
QY 121 CTTACCTTTTGTGATGATATATATATATATATATATATATATATATATATATAT 176
DB 129 CTTACCTTTTGTGATGATATATATATATATATATATATATATATATATATATAT 70
QY 177 CACACATATATTTTACCTGCGACATATATATATATATATATATATATATATATATAT 236
DB 69 CACACATATATTTTACCTGCGACATATATATATATATATATATATATATATATATAT 11
QY 237 ATCTCAGAG 246
DB 10 ATCTCAGAG 1

RESULT 2
LOCUS CT350937 882 bp DNA linear GSS 03-NOV-2005
DEFINITION SUE scrofa genomic clone CH242-413G22, genomic survey sequence.
ACCESSION CT350937
VERSION CT350937.1 GI:79926432
KEYWORDS GSS.
SOURCE SUE scrofa (pig)
ORGANISM SUE scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 882)
AUTHORS Humphrey, S.J., Plumb, R.W. and Durham, J.L.
TITLE Direct Submission


```

/lab host="DH108"
/closure lib="ipCGX14"
/note=Vector: psport1, Site_1: Sal 1, Site_2: Noc 1
Primary Library"

```

Query Match	13.5%	Score 53.8;	DB 4;	length 655;
Best Local Similarity	52.0%;	Pred. NO. 0.0075;		
Matches 118; Conservative	1;	Mismatches 108;	Indels 0;	Gaps 0;

[illegible]

RESULT 5	
CNS03LZU/c	CNS03LZU
LOCUS	868 bp DNA linear
DEFINITION	Tetradon nigroviridis genome survey sequence PUC-ori end of clone 03J04 of library G from Tetradon nigroviridis, genomic survey sequence.

ACCESSION	AL250131	GI:7971143
VERSION	AL250131.1	
KEYWORDS	GSS; genome survey sequence.	
SOURCE	Tetradodon nigroviridis	
ORGANISM	Tetradodon nigroviridis	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphi; Acanthopterygii; Percomorphi; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetradodon.	

REFERENCE	AUTHORS	TITLE	JOURNAL	PUBMED	REFERENCE
1	Roeest Crollius,H., Jallou,O., Daellia,C., Bouneau,L., Fisher,C., Benoit,A., Fitzames,C., Winkler,P., Brottier,P., Quettier,F., Saurin,W. and Weissenbach,J.	Baseline of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence	Nat. Genet. 25 (2), 235-238 (2000)	10835645	2

AUTHORS	TITLE	JOURNAL	PUBMED	REFERENCE
Roast Crollins H., Jallion O., Dastiyar C., Ozouf-Costez C., Pflames C., Fischer C., Bonneau L., Billault A., Quetier F., Saurin M., Bernot A. and Weissenbach J.	Characterization and repeat analysis of the compact genome of the freshwater pufferfish <i>Tetraodon nigroviridis</i>	Genome Res.	10 (7), 939-949 (2000)	
			10899143	
			3 (bases 1 to 868)	

AVRONS TITLE JOURNAL	COMMENT
Genoscope. Direct Submission Submitted (12-APR-2000) Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - FRANCE (E-mail : segreff@genoscope.cns.fr - Web : www.genoscope.cns.fr)	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the <i>Tetradodon nigroviridis</i> genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetradodon .

FEATURES	source	location/Qualifiers
	1..868	/organism="Petrodon nigroviridis"
		/mol_type="genomic DNA"
		/db_xref="taxon:99883"

```

/clone="036J04"
/clone_1jb="g"
/notes=Transcriptome sequence ID : COBG036D02SP1
end : PUC-Orig

```

Query Match	13.5%;	Score 53.8;	DB 14;	Length 868;
Best Local Similarity	76.5%;	Pred. No. 0.0077;		
Matches 62;	Conservative 3;	Mismatches 16;	Indels 0;	Gaps 0;

Dy
Dc
Dd
De

191 AACTGGGACATAAATGTACA 211
:::
75 TAYHCTGTATATATATATA 55

RESULT 6	
FR0018554	
LOCUS	612 bp DNA linear GSS 25-FEB-2004
DEFINITION	F. rubripes GSS sequence, clone 016P03aC3, genomic survey sequence.
ACCESSION	AL011449
VERSION	AL011449.1 GI:2676882
KEYWORDS	GSS; genome survey sequence.
SOURCE	Takifugu rubripes (Fugu rubripes)
ORGANISM	Takifugu rubripes

REFERENCE AUTHORS	TITLE
Elgair, G., Clark, M.S., Meek, S., Smith, S., Warner, S., Edwards, Y.U., Bouchireb, N., Cottage, A., Yeo, G.S., Umrani, Y., Williams, G. and Brenner, S.	Generation and analysis of 25 Mb of genomic DNA from the pufferfish <i>Fugu rubripes</i> by sequence scanning
JOURNAL PUBLISHED	Genome Res. 9 (10), 960-971 (1999)
REFERENCE AUTHORS	2 (bases 1 to 612)
Elgair, G., Clark, M., Smith, S., Meek, S., Warner, S., Umrani, Y.,	

TITLE Williams, G. and Brenner, S.
JOURNAL Direct Submission
COMMENT Submitted (08-DEC-1997) MRC Human Genome Mapping Project Resource Centre Hinxton, Cambridge, CB10 1SB. Email: bhlohe1@hgm.mrc.ac.uk
VECTOR pBluescript II KS
V_type: phagemid
PRIMER: KS
DESCR: One pass dye-terminator sequencing of cosmid cloned genomic

```

FEATURES
sequence.
source
Location/Qualifiers
1..612
/organism="TakiFugu rubripes"
/mol_type="genomic DNA"
/db_xref="taxon:31033"
/clone="016P03a3"
/clone_1tb="cosmid 016P03"

```

[illegible]

Oy CTCGGACATTAATGTACACCAACCATTCCTCAAGTAAACAACTACTCTCAGAAAGTAAAG 252
 Db CNCANTTATATAAGGNAATATNNGCAGTTNAAAGTAAATACCTCTCACTGCAAGAGAAA 251

QY 253 CATATCC 260
 |||||
 DB 252 GATATCC 259

RESULT 7
 DR106610
 LOCUS
 DEFINITION JH118H03L779 Canine cardiovascular system biased cDNA Canis familiaris cDNA similar to Hs cell division cycle associated 2 (CDC42), mRNA sequence.
 ACCESSION DR106610
 VERSION DR106610
 KEYWORDS GI:67565965
 SOURCE
 ORGANISM
 Canis familiaris (dog)
 Canis familiaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE
 1 (bases 1 to 779)
 DISilverstre,D., Yung,C., Gao,Z., Farukhi,Y., Winslow,R.L. and Tomaseilli,G.F.
 Canine cardiovascular system biased cDNA sequences
 Unpublished (2005)
 CONTACT: Gordon F. Tomaseilli
 Johns Hopkins University
 720 Rutland Avenue/Koss 844, Baltimore, MD 21205, USA
 Tel: 4109552774
 Fax: 4105022096
 Email: gtomasei@jhm1.edu.

FEATURES
 source
 Location/Qualifiers
 1..779
 /organism="Canis familiaris"
 /mol_type="mRNA"
 /strain="Mixed"
 /db_xref="taxon:9615"
 /sex="Mixed"
 /clone_lib="Canine cardiovascular system biased cDNA"
 /note="Organ: Mixed; Vector: pCDNA3.1; Site_1: EcoRI; Site_2: XhoI; Adult tissue from eye, lung, aorta, pulmonary artery and brain; neonatal tissue from liver, spleen, thymus, lung, kidney, aorta and brain; 50% is from cardiac tissues."

ORIGIN
 Query Match 13.4%; Score 53.6; DB 9; Length 779;
 Best Local Similarity 67.3%; Pred. No. 0.0086;
 Matches 74; Conservative 1; Mismatches 35; Indels 0; Gaps 0;

QY 122 CTACCTTTTGTACTGTTGATATATATATATACACACACACACACACACAC 181
 |||||
 DB 422 CTCTCTCTCTCTCTCTCTATATATATATATACACACACACACACACACAC 481

QY 182 AATATATTTTAACTCGGACATTAATGTACACCACTTTCAAGTAAATAA 231
 |||||
 DB 482 ACAATATACATATACATACATGATTAATATATATATATATATATATAT 531

RESULT 8
 AQ407117/c
 LOCUS
 DEFINITION HS_5076_A2_H04_SP68 RPCT-11 Human Male BAC Library Homo sapiens genomic clone Plate=652 Col=8 Row=O, genomic survey sequence.
 ACCESSION AQ407117
 VERSION AQ407117.1
 KEYWORDS GI:4429739
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1 (bases 1 to 491)
 Mahatras,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,

TITLE
 JOURNAL
 PUBMED
 COMMENT
 Contact: Mahatras GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCT-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
 Plate: 652 row: O column: 8
 Seq primer: SP6
 Class: BAC ends
 High quality sequence stop: 491.

FEATURES
 source
 Location/Qualifiers
 1..491
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=652 Col=8 Row=O"
 /sex="male"
 /clone_lib="RPCT-11 Human Male BAC library"
 /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"

ORIGIN
 Query Match 13.4%; Score 53.4; DB 11; Length 491;
 Best Local Similarity 74.2%; Pred. No. 0.0093;
 Matches 66; Conservative 1; Mismatches 22; Indels 0; Gaps 0;

QY 136 TCGTTGATATATATATATATATACACACACACACACACACACATTTTAACTC 195
 |||||
 DB 227 TGTAGATATATATATATATATATATACACACACATATACACACACCTTATACACAC 168

QY 196 GGGACATTAATGTACACCACTTATTCAA 224
 |||||
 DB 167 ATATATATATATATACACACGCTTATATA 139

RESULT 9
 B2217953
 LOCUS
 DEFINITION CH230-51817.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
 ACCESSION B2217953
 VERSION B2217953.1
 KEYWORDS GI:23876311
 SOURCE
 ORGANISM
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Rattus.
 REFERENCE
 1 (bases 1 to 600)
 Zhao,S., Shetty,J., Shatman,S., Tsegaye,G., Geer,K., Shwartsbeyn,A., Gebregorgie,E., Overton,L., Russell,D., Chen,D., Riggs,P., de Jong,P. and Fraser,C.M.
 Rat BAC End Sequences from Library CHORI-230 MboI segment Unpublished (1999)
 Other GSSES: CH230-51817.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics

The Institute for Genomic Research
9112 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208

Email: azha@tigr.org
Clones are derived from the rat BAC library CHORI-230
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library
availability, please contact Pieter de Jong (pdjong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or_eirng_information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat_bac_end_intro.html
Plate: 518 row: 1 column: 7
Seq primer: SP6
Class: BAC ends.

FEATURES

Source

Location/Qualifiers
1..600
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SENHed/MCM"
/db_xref="taxon:10116"
/clone="CH230-51817"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note="Vector: PTARBAC1.3; Site 1: Mbol; Site 2: Mbol;
CHORI-230 Rat (BN/SENHed/MCM) BAC library produced by
Pieter de Jong"

ORIGIN

Query Match 13.4%; Score 53.4; DB 11; Length 600;
Best Local Similarity 54.4%; Pred. No. 0.0094;
Matches 105; Conservative 1; Mismatches 87; Indels 0; Gaps 0;
113 AAGTAGGCGCTAAGCTTTTGTGCTGTTGATATATATATATATACACACACACACA 172
169 AAGTAGAGACCTCTCTAGTTTATTTAGGATATATATACACACACACACACACA 228
173 CACACACACATATTTTAACTCGGACATTAATGTACACACCTATTTCAAGTAAAC 232
229 CACACACACACAAAGTATGATCATTTAGCATTTGATATGATTTGTTTATCC 288
233 TACTATCTAGAGATTAAGATATATCTGCTATTTGCTATTAAGTAAAGTCC 292
289 GACACACACAGAAATGACAACTGCTCAAAATACACACACAGCTGTAAGCACCC 348
293 TAACCTTATGCA 305
349 GACCCCAATGGA 361

RESULT 10

LOCUS DW252261 757 bp mRNA linear EST 05-JAN-2006
DEFINITION UT-S-GB0-aad-c-04-0-UI-s1 UT-S-GB0 Euprymna scolopes cDNA clone

ACCESSION DW252261
VERSION DW252261.1 GI:84417660
KEYWORDS EST.

SOURCE Euprymna scolopes
ORGANISM Euprymna scolopes
Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
Decapodiformes; Sepioidae; Sepioidae; Euprymna.

REFERENCE 1 (bases 1 to 757)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery

JOURNAL Genome Res. 6 (9) , 791-806 (1996)
PUBMED 8889548
CONTACT: Soares, MB

COMMENT Coordinated laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250

FEATURES

Source

Location/Qualifiers
1..757
/organism="Euprymna scolopes"
/mol_type="mRNA"
/db_xref="taxon:6613"
/clone="UI-S-GB0-aad-c-04-0-UI"
/issue_type="Euprymna scolopes light Organs"
/dev_stage="0 hours"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-S-GB0"
/note="Vector: pT73-Pac (Pharmacia) with a modified
polylinker; Site 1: EcoR I; Site 2: Not I; UI-S-GB0 is a
cDNA library containing the following tissue(s): Euprymna
scolopes light Organs The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is ATGCAAGAT.
TAG TISSUE=Euprymna_scolopes_light_Organs
TAG_LIB=UI-S-GB0
TAG_SEQ=ATGCAAGAT"

ORIGIN

Query Match 13.3%; Score 53; DB 10; Length 757;
Best Local Similarity 54.2%; Pred. No. 0.012;
Matches 104; Conservative 1; Mismatches 87; Indels 0; Gaps 0;

10 CTTTCTCTATATATTTAGAGTTCTTGAATTAAGCTTAAATTAACCTGTGGCGTACATGA 69
142 CTCATGCAATTAAGTAGGCTGAGACATGACCTTACATATCTCCAAAGAAATATC 201

70 GACTGATGATCGTTAAGCTTGAAGATGATGACAGGAGAGGAGCCACTTT 129
202 TATTTAGGTTAAATTCACAAATTTATTAATGTTTCTTAGATATATATATATAT 261

130 TTGTACTGTTGATATATATATATACACACACACACACACACACATATTT 189
262 ATATATATATATATATATATATATACACACACACACACACACACACAGTA 321

190 TAACTGGGACA 201
322 NAAACTGGGATA 333

RESULT 11

LOCUS CT063242 602 bp DNA linear GSS 01-NOV-2005
DEFINITION Sus scrofa genomic clone PigB-51B11, genomic survey sequence.

ACCESSION CT063242
VERSION CT063242.1 GI:78635936
KEYWORDS GSS.

SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suidae; Suidae; Sus.
 1 (bases 1 to 602)
 REFERENCE
 HUMPHREY, S.J., Plumb, R.W. and Durham, J.L.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (01-NOV-2005) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humphrey@sanger.ac.uk Unpublished
 COMMENT
 This sequence was generated from the T7 end of BAC 51E11. 51E11 is part of the PigBAC BAC Library created by Roslin Institute/RFCGR. Further details: http://www.sanger.ac.uk/Projects/S_scrofa/.
 FEATURES
 source
 1..602
 /organism="Sus scrofa"
 /mol_type="genomic DNA"
 /db_xref="taxon:9823"
 /clone="PigB-51E11"
 /tissue_type="Blood cells"
 /note="Vector pBelobAC11 sex male"
 ORIGIN
 Query Match 13.2%; Score 52.6; DB 14; Length 602;
 Best Local Similarity 67.0%; Pred. No. 0.015; Mismatches 35; Indels 0; Gaps 0;
 Matches 73; Conservative 1; Mismatches 35; Indels 0; Gaps 0;
 78 ATATGCTTACCTCTGGAAGTAGTATGACGGGACGTAGGCTTCTTGTACTG 137
 415 ATATGAGAGACTACTGAAAGTACGAAAGTAGTGTGAGTCTTGGAGGTTTAAATAT 474
 138 TTTAT 186
 475 AT 523
 RESULT 12
 CT376921 768 bp DNA linear GSS 03-NOV-2005
 LOCUS
 DEFINITION
 CT376921
 ACCESSION
 CT376921.1 GI:79989881
 VERSION
 GSS.
 KEYWORDS
 Sus scrofa (pig)
 SOURCE
 Sus scrofa
 ORGANISM
 Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suidae; Suidae; Sus.
 1 (bases 1 to 768)
 REFERENCE
 HUMPHREY, S.J., Plumb, R.W. and Durham, J.L.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (01-NOV-2005) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humphrey@sanger.ac.uk Unpublished
 COMMENT
 This sequence was generated from the SP6 end of BAC 445B16. 445B16 is part of the CHOKI-242 BAC Library created by P. de Jong. Further details: http://www.sanger.ac.uk/Projects/S_scrofa/.
 FEATURES
 source
 1..768
 /organism="Sus scrofa"
 /mol_type="genomic DNA"
 /db_xref="taxon:9823"
 /clone="CH242-445B16"
 /tissue_type="White blood cells"
 /note="Vector pTARBAC1.3_BamHI sex female"
 ORIGIN
 Query Match 13.2%; Score 52.6; DB 14; Length 768;
 Best Local Similarity 63.2%; Pred. No. 0.015; Mismatches 45; Indels 0; Gaps 0;
 Matches 79; Conservative 1; Mismatches 45; Indels 0; Gaps 0;
 142 AT 201

620 AT 679
 202 TAAATGACACCACTATTCAGTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 261
 680 TAT 739
 262 GCTAT 266
 740 TATAT 744
 RESULT 13
 BH116000 796 bp DNA linear GSS 19-JUL-2001
 LOCUS
 DEFINITION
 BH116000
 ACCESSION
 BH116000
 VERSION
 BH116000.1 GI:14955795
 KEYWORDS
 GSS.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.
 1 (bases 1 to 796)
 REFERENCE
 ZHAO, S., Nierman, W., Malek, J., Shateman, S., Akhmet, B., Levins, M., Tesegaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregorj, S., Russell, D., de Jong, P. and Fraser, C.M.
 TITLE
 Mouse BAC End Sequences from Library RPCI-24
 JOURNAL
 Unpublished (1999)
 COMMENT
 Other GSSs: RPCI-24-253C11.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end page: http://www.tigr.org/cdb/bac_ends/mouse/bac_end_intro.html.
 Plate: 253 row: C column: 11
 Seq primer: T7
 Class: BAC ends.
 FEATURES
 source
 1..796
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-24-253C11"
 /sex="Male"
 /cell_type="Spleen/Brain"
 /clone_lib="RPCI-24"
 /note="Vector: pTARBAC1; site 1: BamHI; site 2: BamHI; RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamHI sites using MhoI partially digested male C57BL/6J DNA."
 ORIGIN
 Query Match 13.2%; Score 52.6; DB 11; Length 796;
 Best Local Similarity 65.0%; Pred. No. 0.015; Mismatches 76; Conservative 1; Mismatches 40; Indels 0; Gaps 0;
 Matches 76; Conservative 1; Mismatches 40; Indels 0; Gaps 0;
 118 AGGCTTACCTTTTGTGATATATATATATATATATATATATATATATATATATAT 177
 341 ATGACAT 400
 178 ACACATATATTTTATACCGGAGCATTAATGTACACACACCTATTGAAAGTAAAGTAA 234

[illegible]

Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 7500 Std Error: 0.00
Seq primer: Sp6 ATTAGTGACACTATAG
Class: BAC ends
High quality sequence start: 12
High quality sequence stop: 766.

FEATURES

SOURCE

Location/Qualifiers

1..1015

/organism="Xenopus tropicalis"

/mol_type="genomic DNA"

/db_xref="taxon:8364"

/clone="ISB1-79pi0"

/clone_lib="ISB1"

/note="Vector: pBelBAC11, ISB-1 Xenopus tropicalis BAC Library Segment 1"

ORIGIN

Query Match 13.2%; Score 52.6; DB 12; Length 1015;
Best Local Similarity 57.0%; Pred. No. 0.016; Mismatches 70; Indels 0; Gaps 0;
Matches 94; Conservative 1;

Oy 138 TTGTGATTATATATATACATACACACACACACACACACACACATATTTTAACCTGG 197
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 214 TATTCAGCTTCT 273
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 198 GACATTAATATGTACACCACCTATTCAAGTAAATAAATCTATCTCGAAGTAAAGACATAA 257
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 274 TTCATATTTTTTAAAGCTTAAATAATATTAATCAAACTTCATATGTGATGAAGAAGAAA 333
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 258 TCCTGCTATTTGCTCATTAACGTACGATTAAAGCTGCCTAACCTTAT 302
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 334 CAAATATCAGCAATGTAAATATGAGAAACAATTAATTTCT 378
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 16

CN674247/c

LOCUS CN674247 315 bp mRNA linear EST 17-MAY-2004

DEFINITION A0946A09-5 NIA Mouse Embryonic Stem (ES) cell Lf1+, 48 h, high density) CDNA library (Long) Mus musculus cDNA clone NIA:A0946A09 IMAGE:30771176 5', mRNA sequence.

CN674247

ACCESSION CN674247.1 GI:47440698

VERSION

KEYWORDS

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murine; Mus.

REFERENCE

A1 (bases 1 to 315)
Sharov,A.A., Piao,Y., Matoba,R., Dudekula,D.B., Qian,Y., VanBuren,V., Falco,G., Martin,P.R., Stegg,C.A., Bassey,V.C., Wang,Y., Carter,M.G., Hamatni,T., Alba,K., Akutsu,H., Sharova,L., Tanaka,T.S., Kimber,W.L., Yoshikawa,T., Jaradat,S.A., Pantano,S., Nagaraja,R., Boteler,K.R., Taub,D., Hodas,R.J., Longo,D.L., Nagatsinger,D., Keller,J., Klotz,E., Kelesoe,G., Umezawa,A., Vecovi,A.L., Rossant,J., Kunath,T., Hogan,B.L., Curci,A., D'Urso,M., Kelso,J., Hide,W. and Ko,M.S.
Transcriptome analysis of mouse stem cells and early embryos plus Biol. 1 (3), 410-419 (2003)

JOURNAL COMMENT

Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@gsun-grc.nia.nih.gov
Plate: A0946 row: A column: 09
Seq primer: M13 Reverse
High quality sequence stop: 315
POLYA=NO.

FEATURES

SOURCE

Location/Qualifiers

1..315

/organism="Mus musculus"

/mol_type="mRNA"

ORIGIN

RESULT 17
CN676171/c

ACCESSION	CN676171	GI:47442622
VERSION	CN676171.1	
KEYWORDS	EST.	
SOURCE	Mus musculus	(house mouse)
ORGANISM	Mus musculus	

REFERENCE
AUTHORS

Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciuromorphi; Muroidae; Muridae; Murinae; Mus.
1 (bases 1 to 521)
Sharov, A.A., Piao, Y., Matoba, R., Dudchula, D.B., Qian, Y.,
VanBuren, V., Falco, G., Martin, P.R., Stagg, C.A., Bassey, U.C.,

FEATURES

SOURCE

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="129Sv/EvTac"
/db_xref="nlist:A0976B03-5"
/db_xref="taxon:10090"
/clone="NIA:A0976B03 IMAGE:30774062"
/sex="Male"
/tissue_type="Embryonic Stem Cell"
/cell_line="129.3 ES cells"
/lab_host="DH10B"
/clone_1ib="NIA Mouse Embryonic Stem (ES) cell (Lif+, 48 h, high density) CDNA library (Long)"
/notes="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://gisun.gsc.nia.nih.gov/CDNA) This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMD:11544199]). ES cells were plated at density 3x104/cm2, on gelatin-coated plates and cultured for 48 hrs at 37 OC, 5% CO2. Culture medium: DMEM supplemented with 15% FBS, 2 mM L-glutamine, 0.1 mM NEAA, 1mM Sodium pyruvate, 0.1 mM beta-mercaptoethanol, 1000 U/ml Lif+, 100 U/ml penicillin, and 100 ugc/ml streptomycin. Double-stranded cDNAs were synthesized with an Oligo(dt) primer [Invitrogen: 5'-pGACATGTTCTAGATCGGAGCGGCCCTCTTTTCTTTT-3'] from 2ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker IL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.7 kb. The library was constructed by Yulan Piao."

```

ORIGIN	Query Match	13.1%	Score 52.4;	DB 8;	Length 521;
	Best Local Similarity	59.7%	Pred. No. 0.017;		
	Matches 86;	Conservative 1;	Mismatches 57;	Indels 0;	Gaps 0
Qy	124	ACCTTTTGTCGCTTTGATATATATATATATACATACACACACACACACACACACAY	183		
Db	304	ATCTGTGTGATGTGTGTCATATACATATATATATATATACACACACACACACACACAC	245		
Qy	184	ATATTTTAACTCGGAGATTAATGTACACACACCTATTCAAAAGTAAAAAATACTATCTCAG	243		
Db	244	ATACTATTACCAATGTGCACATATACCAATGCACATTAATAATTAACAGGTTTAA	185		
Qy	244	AAGTAAAGACATTAATCCTGCTATT	267		

Db	184	TTCTAAGAGAGAAACACAGATT	161
RESULT 18			
LOCUS	A0353718	541 bp	DNA
DEFINITION	CITBI-EI-2540J1.TR CITBI-EI Homo sapiens genomic clone 2540J1,		
ACCESSION	A0353718		
VERSION	A0353718.1		
KEYWORDS	GSS.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 541) Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and Venter,J.C.		
TITLE	Use of BAC End Sequences from CalTech Libraries for Sequence-Ready Map Building		
JOURNAL	Unpublished (1997)		
COMMENT	Other GSSes: CITBI-EI-2540J1.TF Contact: Shaying Zhao, William Nierman, Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850 Tel: 301 838 0200 Fax: 301 838 0208 Email: hbe@tigr.org Clones are available from Research Genetics (info@resgen.com). BAC clones and search page: http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html . Seq primer: M13 Reverse Class: BAC ends.		
FEATURES			
source	Location/Qualifiers 1..541 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /clone="2540J1" /sex="male" /cell_type="sperm" /clone_11b="CITBI-EI" /note="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI; Caltech Human BAC Library D"		
ORIGIN			
Query Match	13.1%; Score 52.4; DB 11; Length 541;		
Best Local Similarity	54.7%; Pred. No. 0.017;		
Matches	98; Conservative 2; Mismatches 79; Indels 0; Gaps 0;		
Db	55	GTGGCGTACATGAGACTGATGCTGTTACTCTCGAAGATGAGTATCGACGGGCA	114
Db	308	GTGGATTCATCCATCAAGAAGCTTTGGAGATCTACCCAGATGAGATCGGCTCTCA	367
Db	115	GTGAGGCTTACTCTTTGTAGCTGTTGATATATATATATACATACACACACACACA	174
Db	368	TATATTTCAACGGCATTTATATTTTGGACATATATATATATATATATACACACACA	427
Db	175	CACACACATATTTTAACTCGGGACATAAATGTACACACACTTTCAAAGTAAAACT	233
Db	428	CACACATATATATATATATATATTTTGAAGAACATGTCACACACTTATTTTGAATTAAT	486
RESULT 19			
LOCUS	A0779979	619 bp	DNA
DEFINITION	HS_3113_A1_D12 MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3113 Col=23 Row=G, genomic survey sequence.		
ACCESSION	A0779979		

VERSION	A0779979.1 GI:5682939		
KEYWORDS	GSS.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo		
REFERENCE	1 (bases 1 to 619)		
AUTHORS	Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M. D. and Hood, L.		
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)		
PUBMED	10449764		
COMMENT	Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Clones may be purchased from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu Plate: 3113 row: G column: 23 Seq primer: M13 Reverse Class: BAC ends High quality sequence stop: 619.		
FEATURES	Location/Qualifiers		
Source	1..619		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/clone="Plate=3113 Col=23 Row=G"		
	/sex="male"		
	/clone_lib="CIT Approved Human Genomic Sperm Library D"		
	/note="Organ: sperm; Vector: pBelobAC11; BAC clones in E-Coli DH10B"		
ORIGIN			
Query Match	13.1%; Score 52.4; DB 11; Length 619;		
Best local similarity	56.5%; Pred. No. 0.017; 72; Indels 0; Gaps 0;		
Matches	95; Conservative 1; Mismatches 72; Indels 0; Gaps 0;		
Qy	19 TATATTTACGAGTCTTGAATTAAGCTTAAATACCTGTGGGTACACATGAGACTGATGA 78		
Db	482 TTTAGTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 423		
Qy	79 TACTGTACTCTCTGAAAAGTAGGTATGACGAGGCAAGTGAAGGCTTCTTTTACTGT 138		
Db	422 CACGCGGTACTGTGACCTTAAATATGATGTTGTGACGGTAAATTTCTTTTACCAAAAT 363		
Qy	139 TTGATATATATATATATACATACACACACACACACACACACACACACATTA 186		
Db	362 TAAATATATATATATATATGTTGACACACACACACACACACACACACACACACAC 315		
RESULT 20			
AGS61036	736 bp DNA linear GSS 23-DEC-2004		
LOCUS	Mus musculus molossinus DNA, clone:MSWg01-480G02.T7, genomic survey		
DEFINITION	sequence.		
ACCESSION	AGS61036		
VERSION	AGS61036.1 GI:48321734		
KEYWORDS	GSS.		
SOURCE	Mus musculus molossinus (Japanese wild mouse)		
ORGANISM	Mus musculus molossinus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Mus.		
REFERENCE	1 Abe, K., Noguchi, H., Tagawa, K., Yuzurika, M., Toyoda, A., Kojima, T., Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriwaiki, K. and		

TITLE	Shirotoshi, T.
CONTRIBUTION	Contribution of Asian mouse subspecies <i>Mus musculus molossinus</i> to genomic constitution of strain C57BL/6J, as defined by BAC-end sequence-SNP analysis
JOURNAL	Genome Res. 14 (12), 2439-2447 (2004)
PUBMED	15574823
REFERENCE	2 (bases 1 to 736)
AUTHORS	Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (11-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Shiho-chou, Tsukuba, Ibaraki, 305-0857, Japan
COMMENT	(E-mail: hattori@sc.riken.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170) Clones are derived from the mouse BAC library MSNg01. For BAC library availability, please contact Kunihisa Abe (abe@tc.riken.jp), Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan phone: 81-298-36-9189, fax: 81-298-36-9199 e-mail: abe@tc.riken.jp
PRIMERS	Sequencing : T7
LIBRARY	Vector : pBACe3.6
R.Site 1	: EcoRI
R.Site 2	: EcoRI
FEATURES	Location/Qualifiers 1..736 /organism="Mus musculus molossinus" /mol_type="genomic DNA" /sub_species="molossinus" /db_xref="taxon:57486" /clone="MSNg01-480G02.T7" /sex="male" /tissue_type="mixture of kidney and spleen" /clone_id="MSNg01 Mouse Male BAC Library"
ORIGIN	Query Match Best Local Similarity 13.1%; Score 52.4; DB 14; Length 736; Matches 86; Conservative 1; Mismatches 57; Indels 0; Gaps 0; 142 ATATATATATATACATACACACACACACACACACACATATTTAACTGGGACA 201 277 ACACATACATACACACACATACATACACACACACACACACCTCTATATTTTCA 336 202 TAAATGTACACACCTATTCAAAGTAAACTATCTCAGAGTAAAGCATATCTT 261 337 TATATATATATAGATCACTTAATATGTAAAGTGTCAATTAATCGAAGAAACATATCAT 396 262 GCTATGTGCATTAACGTAGCAGTTA 285 397 CCTAAATACCTATAGCTTCATATA 420
RESULT 21	AKI42564
LOCUS	AKI42564/c
DEFINITION	Mus musculus 0 day neonate kidney cDNA, RIKEN full-length enriched library, clone:D630038012 product:unclassified, full insert sequence.
ACCESSION	AKI42564
VERSION	AKI42564.1 GI:74216314
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.
REFERENCE	1 Carninci, P. and Hayashizaki, Y.
AUTHORS	High-efficiency full-length cDNA cloning
TITLE	

JOURNAL	Meth. Enzymol.	303,	19-44	(1999)
PUBMED	10349636			
REFERENCE				
AUTHORS				
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoh, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res.	10	(10),	1617-1630 (2000)
PUBMED	11042159			
REFERENCE				
AUTHORS				
TITLE	Shibata, K., Itoh, M., Aizawa, K., Nagaoke, S., Sasaki, N., Carninci, P., Komoh, H., Atiyama, J., Niehl, K., Kitsumai, T., Tashtir, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsuno, H., Sakaguchi, S., Ikegami, T., Kawhiwa, K., Fujiwara, S., Inoue, K., Togaw, Y., Izawa, M., Ohara, E., Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Ozeaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.			
JOURNAL	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer			
PUBMED	Genome Res.	10	(11),	1757-1771 (2000)
REFERENCE	11076861			
AUTHORS				
Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Komoh, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Klyosawa, H., Kondoh, S., Yamanka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Pleischmann, M., Gaasterland, T., Gissi, C., Kling, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Stambli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C., Fletcher, C., Fujita, M., Gariboldi, M., Guscinich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mommaerts, P., Nordone, P., Ring, B., Ringswald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyokawa, K., Wang, K.H., Weiler, C., Whitaker, C., Wilting, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohseki, S. and Hayashizaki, Y.				
CONSRMT	RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium			
TITLE	Functional annotation of a full-length mouse cDNA collection			
JOURNAL	Nature	409	(6821),	685-690 (2001)
PUBMED	11217851			
REFERENCE				
AUTHORS				
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanka, I., Klyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bull, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanpin, A., Matsuda, H., Batool, S., Beisel, K.W., Blake, J.A., Brad, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dreagan, T.A., Fletcher, C.F., Forrest, A., Fraser, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gong, J., Grimmond, S., Guscinich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedziarski, R.M., King, B.L., Konaga, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pereira, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B., Ringswald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, K., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilting, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatune, N., Hirozane-Kishikawa, T., Komi, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawaji, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashitume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,				

Plate: 3237 row: F column: 24
Class: BAC ends
High quality sequence stop: 494.

```

1..494
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="plate:3237 COL-24 Row=F"
/sex="male"
/clone_lib="CTF Approved Human Genomic Sperm Library D"
note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
-D-Col1 DH10B"

```

ORIGIN

Query Match	13.1%;	Score 52.2;	DB 11;	Length 494;
Best Local Similarity	51.1%;	Pred. No. 0.019;		
Matches 120;	Conservative 1;	Mismatches 114;	Indels 0;	Gaps 0;

Q7	146	ATATATATATACATACACACACACACACACACATATTTAACTCGGACATPAA	205
D8	63	TACTGTGGAAAATATTTATATGAGGCAATTGTATATATATATATATATATATGT	122
Q7	86	TACCTCTGGAAAGTAGATGTGACGGGCAAGTGAAGCCATACCTTTTGGATCGTTGGAT	145
D8	243	CTCTAAAGTATTAACGCTTAACTCTAAATAATATATAGATTAACAAGTACTTTC	297
Q7	206	TGTACACCACTTTCAAAGTAAAACTATCTCAGAGTAAAGACATATCTTGCTA	265
D8	183	TTATGTTTATATTTGAAATTAATAAACTTTTCAATAGAAAAAACTPAAAAATGACAG	242
Q7	266	TTGTCTATTAACATGACAGTTAAAGCTGCTAACCTTTATGCACTTCAGTTCCCTAC	320
D8	243	CTCTAAAGTATTAACGCTTAACTCTAAATAATATATAGATTAACAAGTACTTTC	297

RESULT	23
BQ188992/c	
LOCUS	BQ188992 660 bp mRNA linear EST 30-APR-2007
DEFINITION	UI-E-EU1-a]s-h-14-0-UI r1 UI-E-EU1 Homo sapiens cDNA clone UI-E-EU1-a]s-h-14-0-UI 5' , mRNA sequence.

REFERENCE	1 (bases 1 to 660)
AUTHORS	Ronaldo,M.F., Lennon,G. and Soares,M.B.
TITLE	Normalization and subtraction: two approaches to facilitate gene
JOURNAL	Genome Res. 6 (9), 791-806 (1996)
PubMed	8889548
COMMENT	Contact: Soares, MB

FEATURES
source

Location/Qualifiers
1. .660

ORIGIN

Query Match	13.1%;	Score 52.2;	DB 3;	Length 660;
Best Local Similarity	72.5%;	Pred. No. 0.019;		
Matches 66;	Conservative 1;	Mismatches 24;	Indels 0;	Gaps 0;

[illegible]

RESULT	24
BQ184399	
LOCUS	BQ184399
DEFINITION	BQ184399 665 bp mRNA linear EST 30-APR-2007
ACCESSION	U1-E-BJ1-a.js-h-14-0-U1.s1 U1-E-BJ1 Homo sapiens cDNA clone
VERSION	U1-E-BJ1-a.js-h-14-0-U1 3', mRNA sequence.
KEYWORDS	BQ184399
SOURCE	BQ184399.1 GI:20359950
ORGANISM	EST.
	Homo sapiens (human)
	Homo sapiens

REFERENCE	1	(bases 1 to 665)
AUTHORS	Bonaldo,M.P., Lennon,G. and Soares,M.B.	
TITLE	Normalization and subtraction: two approaches to facilitate discovery	
JOURNAL	Genome Res.	6 (3), 791-806 (1996)
PUBMED	889548	
COMMENT	Contact: Soares, MB	

The following repetitive elements were found in this cDNA sequence: 263-500, >LINE33#LINE/L1 (matched complement) 334-500, >LINE7#LINE/L1 (matched complement) 513-583, >(TA)n#Simple repeat (matched complement)
Seq primer: M13 REVERSE.

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 The following repetitive elements were found in this cDNA
 sequence: 74-144, >(TA)n#simple_repeat (matched complement)
 157-394, >LIMB3#LINE/L1 157-333, >LIMB7#LINE/L1
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

Location/Qualifiers
 1..665
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-B-EJ1-a15-h-14-0-UI"
 /issue_type="fetal eyes, lens, eye anterior segment,
 optic nerve, retina, Retina Foveal and Macular, RPE and
 Choroid"
 /dev_stage="fetal and adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-B-EJ1"
 /note="Organ: eye; Vector: pUT73-Pac (Pharmacia) with a
 modified polylinker; Site_1: Ecor I; Site_2: Not I;
 UI-B-EJ1 is a subcloned cDNA library constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an Ecor I adaptor, digested
 with Not I, and cloned directionally into pUT73-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tags for this library are: fetal eyes,
 AGAATCAGCA; lens, CGATTAGCGA; eye anterior segment,
 AATGCCGCGAT; optic nerve, CCATTAGTG; retina, CCGCG; Retina
 Foveal and Macular, GTCC; RPE and Choroid, ACTTA. This
 library was created for the program, Gene Discovery in the
 Visual System, supported by National Eye Institute (NEI).
 TAG TISSUE=Foveal and Macular Retina
 TAG LIB=UI-B-EJ1
 TAG_SEQ=GTCC"

ORIGIN

Query Match 13.1%; Score 52.2; DB 3; Length 665;
 Best Local Similarity 72.5%; Pred. No. 0.019;
 Matches 66; Conservative 1; Mismatches 24; Indels 0; Gaps 0;

115 GTGAGGCTACCTTTTGTACTGTTGATATATATATATACACACACACACA 174
 |||||
 67 GGGAGGATATATATATATATATATATATATATATATATACACACACACA 126
 |||||
 175 CACACACATATATTTAACTGGGACATATA 205
 |||||
 127 CACACACATATATATATATATATATATATATATATATATACACA 157
 |||||

RESULT 25
 DW264376 820 bp mRNA linear EST 05-JAN-2006
 LOCUS
 DEFINITION UI-S-GN1-abn-d-03-0-UI s1 UI-S-GN1 Euprymna scolopes cDNA clone
 ACCESSION UI-S-GN1-abn-d-03-0-UI 3, mRNA sequence.
 VERSION DW264376
 KEYWORDS DW264376.1 GI:84429779
 EST.

SOURCE

ORGANISM Euprymna scolopes
 Euprymna scolopes
 Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
 Decapodiformes; Sepioidae; Sepioidae; Euprymna.

REFERENCE

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 PUBMED 8889548

COMMENT

Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. M. J. McFall-Ngai, University of
 Wisconsin-Madison
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/squid.html
 The following repetitive elements were found in this cDNA
 sequence: 306-354, >(CA)n#simple_repeat (matched complement)
 518-546, >AT-rich#low_complexity (matched complement)
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

Location/Qualifiers
 1..820
 /organism="Euprymna scolopes"
 /mol_type="mRNA"
 /db_xref="taxon:6613"
 /clone="UI-S-GN1-abn-d-03-0-UI"
 /issue_type="Euprymna scolopes 12-hour symbiotic"
 /dev_stage="12 hours symbiotic"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-S-GN1"
 /note="Vector: pUT73-Pac (Pharmacia) with a modified
 polylinker; Site_1: Ecor I; Site_2: Not I; UI-S-GN1 is a
 normalized library derived from UI-S-GN1. The library was
 constructed and normalized according to Bonaldo, Lennon
 and Soares, Genome Research, 6:791-806, 1996. First strand
 cDNA synthesis was primed with an oligo-dT primer
 containing a Not I site. Double stranded cDNA was ligated
 to an Ecor I adaptor, digested with Not I, and cloned
 directionally into pUT73-Pac vector. The oligonucleotide
 used to prime the synthesis of first strand cDNA contains
 a library tag sequence that is located between the Not I
 site and the (dT)18 tail. The sequence tag for this
 library is GAGCCTGACT
 TAG TISSUE=Euprymna_scolopes_12-hour_symbiotic
 TAG LIB=UI-S-GN1
 TAG_SEQ=GAGCCTGACT"

ORIGIN

Query Match 13.1%; Score 52.2; DB 10; Length 820;
 Best Local Similarity 65.2%; Pred. No. 0.019;
 Matches 75; Conservative 1; Mismatches 39; Indels 0; Gaps 0;

96 AAGTAGGTATGACAGGAGGAGGCTACCTTTTGTACTGTTGATATATATATAC 155
 |||||
 396 ATGGGCTATCTTTTGTAAATATTTCTCTCTTTTCTTTTAAATATATATATAC 337
 |||||
 156 ATACACACACACACACACACACATATTTAACTGGGACATATAATGTAC 210
 |||||
 336 ACACACACACACACACACACACACACACATTTTGAAGCAAAAAAGCTTC 282
 |||||

RESULT 26

DU046739 950 bp DNA linear GSS 12-AUG-2005
 LOCUS
 DEFINITION clone LB_HB80036P17 5, genomic survey sequence.
 ACCESSION DU046739
 VERSION DU046739.1 GI:72473946
 KEYWORDS GSS.

SOURCE

ORGANISM Lycopersicon esculentum (Solanum lycopersicum)
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.

[illegible]


```

1..501
/organism="Apis mellifera"
/mol_type="mRNA"
/strain="mixed strains of European bees, predominantly
A.m. ligustica"
/db_xref="taxon:7460"
/clone="IB170022A20D02"
/sex="female"
/tissue_type="brain"
/dev_stage="adult worker honey bee"
/lab_host="MDH08"
/clone_lib="Bee Brain Normalized/Subtracted Library, BB17
note: Organ: Brain Vector: pRT3D-Pac1, Site_1: 5'ORF;
site_2: Nc1, This BB17 cDNA library was generated by

```

```

FEATURES
    source
        Location/Qualifiers
            1..649
                /organism="Bos taurus"
                /mol_type="genomic DNA"
                /strain="Holstein"
                /db_xref="taxon:9913"
                /clone="RPC142_143F5"
                /sex="male"
                /cell_type="blood"
                /clone_id="RPC1-42"
                /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
                RPC1-42 Bovine BAC library (Male) produced by Pieter de
                Jong"
    ORIGIN

```

	Query Match	13.0%	Score 51.8;	DB 13;	Length 649;	
	Best Local Similarity	58.9%;	Pred. No. 0.024;	Mismatches 62;	Indels 0;	Gaps 0;
	Matches	89;	Conservative	0;		
Dy	131 TGTACTGTTTGAATATATATATATACATACACACACACACACACATATATT	190				
Dd	157 TGGAGAGATC	216				
Dy	191 AACTCGGACATTAATGTACACACCCTTTCAAAGTAAAACTACTATCTGAGAAGTAA	250				
Dd	217 ATTTTGCTGTCACTCTGGAAGTAGACACTTTTAAATAAATAAATTAAGACTTAATGAA	276				
Dy	251 GACATTAATCTGCTATTGTGCATTAACCTAGCA	281				
Dd	277 ACCCTTAAATCTATTCTTAACCTACTACCTTACA	307				
	RESULT 33					
	CB940280/c					
	LOCUS					
	DEFINITION	CB940280	528 bp	mRNA	linear	EST 29-APR-2003
	IPCGX14_16_P10_23 IPCGX14 Ictalurus punctatus cDNA clone					
	IPCGX14_16_P10_5', mRNA sequence.					
	ACCESSION	CB940280				
	VERSION	CB940280.1	GI:30225684			
	KEYWORDS	EST.				
	SOURCE	Ictalurus punctatus (channel catfish)				
	ORGANISM	Ictalurus punctatus				
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes; Ictaluridae; Ictalurus. 1 (bases 1 to 528) Jensen,K.T., Marx,G., Miller,N., Clem,L.M. and Wilson,M. Identification of expressed genes in mixed leukocyte culture and in a macrophage cell line in channel catfish, Ictalurus punctatus unpublished (2003) Contact: Waldbieser GC Catfish Genetics Research Unit USA-Agricultural Research Service 141 Experimental Station Road, Stoneville, MS 38776, USA Tel.: 662 686 3593 Fax: 662 686 3567 Email: gwaldbieser@ars.usda.gov Single pass sequencing. Bases called with Phred v0.000925.c. Low quality bases and vector trimmed with Lucy vl.16. Plate: 16 row: F Column: 10 Seq primer: T7. Location/Qualifiers				
	FEATURES					
	source	1..528				
	/organism="Ictalurus punctatus"					
	/mol_type="mRNA"					
	/strain="commercial"					
	/db_xref="taxon:7998"					
	/clone="IPCGX14_16_P10"					
	/cell_type="macrophage"					
	/cell_line="427A"					
	/dev_stage="adult"					
	/lab_host="DH10B"					
	/clone_idb="IPCGX14"					
	/note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I; Primary library"					
	ORIGIN					
Dy	Query Match	12.9%;	Score 51.6;	DB 4;	Length 528;	
Dd	Best Local Similarity	64.7%;	Pred. No. 0.027;	Mismatches 40;	Indels 0;	Gaps 0;
	Matches	75;	Conservative	1;		
Dy	121 CCTACCTTTTTTACTGTTGATATATATATATATATATACATACACACACACACACA	180				
Dd	477 CATCATACATACACACACACA	418				
Dy	181 CAYATATTTTAACTCGGACATAAATGTACACCACCTTATCAAAGTAAAATACTACT	236				

```

Db      417 CACACCTTATTTTTCATATCTTTTGATACCGCTTTGCCAGACAAGTACT 362
||:|||||
RESULT 34
BX162787/c      694 bp      DNA      linear      GSS 13-MAR-2003
LOCUS
DEFINITION      Dario rerio genomic clone DKEX-124H4, genomic survey sequence.
ACCESSION      BX162787
VERSION      BX162787.1
KEYWORDS      GSS.
SOURCE      Dario rerio (zebrafish)
ORGANISM      Dario rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Osteichthyes;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 694)
Humphray,S.J., Huckle,E. and Durham,J.L.
Direct Submission
Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humphray@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 124H4. 124H4 is
part of the Daniokey BAC library created by R. Plasterk and N.V.
Keijne. Further details:
http://www.sanger.ac.uk/Projects/D_rerio/.
Location/Qualifiers
1..694
/organism="Dario rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEX-124H4"
/issue_type="Testis"
/notice="vector pindigoBAC-536"

ORIGIN
Query Match      12.9%; Score 51.6; DB 14; Length 694;
Best local similarity 66.7%; Pred. No. 0.027;
Matches 72; Conservative 1; Mismatches 35; Indels 0; Gaps 0;

Qy      183 YATATTTAACTCGGACATAAATGTAACCACTTATTCAAAGTAAA 230
||:|||||
Db      591 TACATACATACATATATATATATATTAAGATACATAAAGTAAA 544

RESULT 35
BX128732/c      827 bp      DNA      linear      GSS 13-MAR-2003
LOCUS
DEFINITION      Dario rerio genomic clone DKEX-275J9, genomic survey sequence.
ACCESSION      BX128732
VERSION      BX128732.1
KEYWORDS      GSS.
SOURCE      Dario rerio (zebrafish)
ORGANISM      Dario rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Osteichthyes;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 827)
Humphray,S.J., Huckle,E. and Durham,J.L.
Direct Submission
Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humphray@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 275J9. 275J9 is
part of the Daniokey BAC library created by R. Plasterk and N.V.
Keijne. Further details:
http://www.sanger.ac.uk/Projects/D_rerio/.
Location/Qualifiers
1..827

FEATURES
source

```


	Matches	82;	Conservative	1;	Mismatches	52;	Indels	0;	Gaps	0;
OY	136	TGTTTGATATATATATATATACATACACACACACACACACACACATATTTAACTC	195							
Db	191	TGATTTTAAAAATCATCACACACACACACACACACACACACACACACACACAC	250							
OY	196	GGGACATTAATGTATACACCACTATTCCAAGTAAAACTATCTCAGAGTAAAGACAT	255							
Db	251	CACACTGATGGTTGGTCTTAACCTTAGACGACACACATATCTCTTCCAAATTAAGACAT	310							
OY	256	AATCTGCTATTTGTC	270							
Db	311	AGTATTTCATTTGGC	325							

RESULT	38		
AZ514487			
LOCUS			
DEFINITION	AZ514487	178 bp	DNA linear GSS 05-OCT-2000
ACCESSION	U0361P24	Mouse 10kb plasmid U03C1M library Mus musculus genomic	
VERSION	clone U03C1M0361P24 F,	genomic survey sequence.	
KEYWORDS	AZ514487	GI:10695899	
SOURCE	GSS.		
ORGANISM	Mus musculus (house mouse)		
	Mus musculus		

REFERENCE AUTHORS

TITLE	Mouse whole genome scaffolding with paired end reads from 10kb
JOURNAL	plasmid inserts
COMMENT	unpublished (2000)
	Contact: Robert B. Weiss

Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0361 row: F column: 24
 Seq primer: CATTGTAACGACGCGCACT
 Class: plasmid ends
 High quality sequence stop: 178.

FEATURES	SOURCE
----------	--------

/organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UDGCM10361F24"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UDGCM library"
 /note="Vector: pMD24inv. Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (<http://www.jax.org/resources/documents/dnares/>). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repeated with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptered DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pMD2 (gi|4732111.gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated

ORIGIN

Query Match	12.8%;	Score 51.2;	DB 11;	Length 178;
Best Local Similarity	63.1%;	Pred. No. 0.031;		
Matches 77;	Conservative 1;	Mismatches 44;	Indels 0;	Gaps 0;

with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

[illegible]

RESULT 39	541 bp	DNA	linear	GSS 24-JAN-2004
AZ703674/c	AZ703674			
LOCUS	RPCT-23-226G23.TV	RPCT-23	Mus musculus genomic clone	
DEFINITION	RPCT-23-226G23, genomic survey sequence.			
ACCESSION	AZ703674			
VERSION	AZ703674.1	GI:12428297		
KEYWORDS	GSS.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			

REFERENCE AUTHORS

TITLE Mouse BAC End Sequences from Library RPCI-23
JOURNAL Unpublished (1999)
COMMENT Contact: Shaying Zhao

Zhao, S., Nierman, W., Feldman, D., Malek, J., Shachar, S., de
Akin, B., Levins, M., McGann, S., Teegate, G., Geer, K., Krol, M.,
Jong, P. and Fraser, C. M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@mai.cho.org). Clones may be purchased from BACPAC
Resources (<http://www.choi.org/bacpac/orderinfoframe.htm>). BAC end
page: http://www.tigr.org/db/bac_ends/mouse/bac_end_intro.html
Plate: 226 Row: G Column: 23
Seq primer: T7
Class: BAC ends

FEATURES	Location/Qualifiers
Source	1. .541

```

1. .541
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCT-23-226G23"
/sex="Female"
/lab_host="DH10B"
/clone_jib="RPCT-23"
/note="Organ: Kidney/Brain; Vector: pBac3.6; Site_1:
EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBac3.6 vector at the

```

Clemson University Genomics Institute
Clemson University

100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: twing@clemson.edu
 Seq primer: TAAATACGACTCCTACTATAGGG
 Class: BAC ends
 High quality sequence stop: 376.
 Location/Qualifiers

FEATURES

1.607
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="japonica"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="nbx0028m24f"
 /tissue_type="leaf"
 /lab_host="E. coli DH10B"
 /clone_lib="CUGI Rice BAC Library"
 /note="Vector: pBeloBAC11; Site 1: HindIII; Site 2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9%. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

ORIGIN

Query Match 12.8%; Score 51; DB 11; Length 607;
 Best Local Similarity 58.4%; Pred. No. 0.038;

Matches 87; Conservative 1; Mismatches 61; Indels 0; Gaps 0;

QY 123 TACCTTTTGAAGCTTGTATATATATATATACACACACACACACACA 182
 DB 407 TATATATATATATATATATATATATATATATATATACACACACACACA 348
 QY 183 YATATTTTAACTCGGACATTAATGTACACCACTATTCAAAGTAAACTATCTATCTCA 242
 DB 347 CACACTAGCAAAATGCCGCTGCTGCGTACGGGTGTTTAAAACTAAATTTGTCAA 288
 QY 243 GAAGTAAAGACATATCTCTGATTTGCA 271
 DB 287 TAAAAAATAATAAATAGTGTGTTTAA 259

RESULT 43
 DN877029 652 bp mRNA linear EST 21-APR-2005
 LOCUS nae10g01.x1 Dog eye eye minus lens and cornea. Unnormalized (nae)
 DEFINITION Canis familiaris cDNA clone nae10g01.3', mRNA sequence.
 ACCESSION DN877029
 VERSION DN877029.1 GI:62846984
 KEYWORDS EST.
 ORGANISM Canis familiaris (dog)
 Canis familiaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 Canis.
 1 (bases 1 to 652)
 Wistow, G.
 NIBBank analysis of Dog eye minus lens and cornea
 JOURNAL Unpublished (2005)

COMMENT

Contact: Wistow G
 Section on Molecular Structure and Function
 National Eye Institute
 6/331, NIH, Bethesda, MD 20892-2740, USA
 Tel: 301 402 3452
 Fax: 301 496 0078
 Email: gwaem@helix.nih.gov
 plate: 10 row: 9 column: 01
 Seq primer: Universal -21M13 Forward.
 Location/Qualifiers

FEATURES

source

1.652

/organism="Canis familiaris"
 /mol_type="mRNA"
 /strain="Beagle"
 /db_xref="taxon:9615"
 /clone="nae10g01"
 /tissue_type="eye minus lens and cornea"
 /dev_stage="Adult"
 /lab_host="EMDH10B"
 /clone_lib="Dog eye eye minus lens and cornea."
 /note="Unnormalized (nae)"
 /note="Organ: Eye; Vector: pCMVSPORT6; RNA was extracted from dog eye minus lens and cornea tissue. A directionally cloned cDNA library in the pCMVSPORT6 vector (Invitrogen) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System, full details of which are contained in the manufacturer's instruction manual (http://www.lifetechn.com/). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGACTGTCTTGAATGCGAGCGCGCC(T15-3')]. cDNA was cloned in Not I/Sal I sites. EST analysis was performed at the NIH Intramural Sequencing Center (NISC). Analyzed data available through http://neibank.nhl.nih.gov."

ORIGIN

Query Match 12.8%; Score 51; DB 9; Length 652;
 Best Local Similarity 56.6%; Pred. No. 0.038;

Matches 90; Conservative 2; Mismatches 67; Indels 0; Gaps 0;

QY 68 GAAGCTGATGATCTGTTACTCTGTGGAAGTAGTAGGCAAGTAGGCTTACT 127
 DB 204 GATACCTGTTTATTTTATTCCTCTGGAATTAACAAAGCTTTCTAATTTGTGAC 263
 QY 128 TTTTGTCTGTTGATATATATATATATATACACACACACACACACATAT 187
 DB 264 ATATATATGTTATATATATATATATATATATATATATACACACACACATATAT 323
 QY 188 TTTTACTCGGACATTAATGTACACCACTATTCAAAGT 226
 DB 324 GTCAAGAATATATATATATTTTACTCAAAATTTTACT 362

RESULT 44
 AO923572 672 bp DNA linear GSS 21-DEC-1999
 LOCUS RPCI-23-299K23.TV RPCI-23 Mus musculus genomic clone
 DEFINITION RPCI-23-299K23, genomic survey sequence.
 ACCESSION AO923572
 VERSION AO923572.1 GI:6612575
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Muridae; Murine; Mus.
 1 (bases 1 to 672)
 Zhao, S., Niernan, W., Feldblum, T., Malek, J., Shatsman, S.,
 Aktiret, B., Levins, M., McGinn, S., Tsegaye, G., Geer, K., Krol, M., de
 Jong, P. and Fraser, C.M.
 Mouse BAC End Sequences from Library RPCI-23
 JOURNAL Unpublished (1999)
 COMMENT Contact: Shaying Zhao

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: sznaco@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tcd/bac_ends/mouse/bac_end_intro.html
Plate: 299 row: K column: 23
Seq primer: 77

Class: BAC ends.

FEATURES

source

Location/Qualifiers
1..672
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-299K23"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methyase. Size selected DNA was cloned into the pBAC3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

ORIGIN

Query Match 12.8%; Score 51; DB 11; Length 672;
Best Local Similarity 59.6%; Pred. No. 0.038;
Matches 84; Conservative 1; Mismatches 56; Indels 0; Gaps 0;
QY 78 ATATCTGTTACTCTGAAAGTAGATGCGAGGCGCTTCTTTGACTG 137
DB 532 ATATGTGAGACCCCTCTCAAAAATAGACATTAGAAATATCATCTATTATATATTC 591
QY 138 TTGAT 197
DB 592 ATAGATTTTATATATATATATATATATATATATATATATATATAT 651
QY 198 GACATTAATGTACACCACTTA 218
DB 652 TAAAACTGCTACTAGACTTA 672

RESULT 45
CT398920/c 835 bp DNA linear GSS 03-NOV-2005
LOCUS
DEFINITION Sus scrofa genomic clone CH242-402G15, genomic survey sequence.
ACCESSION CT398920
VERSION CT398920.1 GI:80050005
KEYWORDS GSS.
SOURCE
ORGANISM Sus scrofa (pig)
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
Sus.

REFERENCE
1 (bases 1 to 835)
Humphray, S.J., Plumb, R.W. and Durham, J.L.
Direct Submission
Submitted (01-NOV-2005) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humphray@sanger.ac.uk Unpublished
details: http://www.sanger.ac.uk/Projects/S_sus/

COMMENT
This sequence was generated from the SPE end of BAC 402G15. 402G15
is part of the CHORI-242 BAC library created by P. de Jong. Further
details: http://www.sanger.ac.uk/Projects/S_sus/

FEATURES

Location/Qualifiers

source

1..835
/organism="Sus scrofa"
/mol_type="genomic DNA"
/db_xref="taxon:9823"
/clone="CH242-402G15"
/clone_lib="white blood cells"
/note="vector pTRBAC1.3_BamHI
sex female"

ORIGIN

Query Match 12.8%; Score 51; DB 14; Length 835;
Best Local Similarity 59.6%; Pred. No. 0.039;
Matches 84; Conservative 1; Mismatches 56; Indels 0; Gaps 0;
QY 97 AGTAGGTATGCGAGGCAAGTACCTTTTGTGACTGTTGATATATATATACA 156
DB 757 ACTGTGATGAGACATGATGAGAGAGATGAGTAATGTGTATATATATACACACA 698
QY 157 TACACACACACACACACACACACACACATATATATATATATATATATATATATACACACC 216
DB 697 CACACACACACACACACACACACACATATATATATATATATATATATATATATATATATAT 638
QY 217 TATTCAAAGTAAACTACTTA 237
DB 637 AATTGACATTCATTTATAGTA 617

RESULT 46
DW594541 1072 bp mRNA linear EST 18-JAN-2006
LOCUS
DEFINITION CGX113-F04.5 y1d-8 SHGC-CGX2 Gasterosteus aculeatus cDNA clone
ACCESSION CGX113-F04.5, mRNA sequence.
VERSION DW594541
KEYWORDS DW594541.1 GI:85167999
SOURCE EST.
ORGANISM Gasterosteus aculeatus (three spined stickleback)
Gasterosteus aculeatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
Gasterosteidae; Gasterosteus.
1 (bases 1 to 1072)
Kingsley, D.M., Peichel, C., Knecht, A., Balabhadra, S., Grimwood, J.,
Dickson, M., Schmutz, J., and Myers, R.W.
Expressed sequence tags from Gasterosteus aculeatus (2004)
Unpublished (2004)
Contact: Grimwood, Jane
Stanford Human Genome Center
Stanford University School of Medicine
975 S California Ave, Palo Alto, CA 94304, USA
Tel: 650 320 5917
Fax: 650 320 5801
Email: jane@hgsc.stanford.edu
Plate: 113

High quality sequence stop: 855.

FEATURES

source

Location/Qualifiers
1..1072
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Conner Creek sticklebacks, WA USA"
/db_xref="taxon:69293"
/clone="CGX113-F04"
/sex="mixed male and female"
/tissue_type="eyes"
/dev_stage="adult"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="SHGC-CGX2"
/note="Vector: Express 1; Total and poly A+ RNA was isolated from the indicated stickleback tissue, and a cDNA library was constructed in the Express 1 plasmid vector by Open Biosystems. First strand cDNA synthesis was primed with an 54 bp linker primer containing an oligodT sequence preceded by a synthetic NotI site (first strand primer:

5'-GACTAGTTCTAGATCGGAGCGGCCCC(T)25-3'). Following second strand synthesis, cDNAs were made blunt at the end corresponding to the original 5' prime end of mRNA, and cloned directionally into the NotI and EcoRV sites of Express 1. Note that the EcoRV site is typically destroyed in the blunt end cloning, leaving a junction of the form 'xxxxATC' (where is ATC is the second half of the EcoRV site, and xxx is derived from the cDNA sequence). A map of the Express 1 vector is available at: http://www.openbiosystems.com/cdna_library_construction_fa_q.php#8 The primary library was transformed and amplified in DH10B (T1 phage resistant) bacteria."

ORIGIN

Query Match 12.8%; Score 51; DB 10; Length 1072;
Best Local Similarity 56.4%; Pred. No. 0.04; Mismatches 71; Indels 0; Gaps 0;
Matches 93; Conservative 1;

123 TACCTTTTGTACTGTTGATATATATATATATACATACACACACACACACACA 182
120 TAT 179
183 YATATTTTAACTGGGACATTAATGTACACCACTATTCAAAGTAAACTACTATCTCA 242
180 CATATTTATATATATATATCTCAAAATACATAGTCTCAGCAGCAAGAAAACGTCACATTCC 239
243 GAGTAAGACATTAATCTGCTATGTCATTAATCTAGACATTAAG 287
240 TTGATTAACCAAGAAAGTAAATATCTTCAATTAATAAAAAAAG 284

RESULT 47

DN738453 1381 bp mRNA linear EST 31-MAR-2005
LOCUS CNB98-F11.x1d-c SHGC-CNB Gasterosteus aculeatus cDNA clone
ACCESSION DN738453
VERSION DN738453.1 GI:62115463
KEYWORDS EST.

SOURCE Gasterosteus aculeatus (three spined stickleback)
ORGANISM Gasterosteus aculeatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
Gasterosteidae; Gasterosteus.

1 (bases 1 to 1381)

REFERENCE 1 Kingsley, D.M., Peichel, C., Balabhadra, S., Grimwood, J., Dickson, M.,
Schmutz, J., and Myers, R.M.
Expressed sequence tags from *Gasterosteus aculeatus*
Unpublished (2003)
Contact: Grimwood, Jane
Stanford Human Genome Center
Stanford University School of Medicine
975 S California Ave, Palo Alto, CA 94304, USA
Tel: 650 320 5917
Fax: 650 320 5801
Email: jane@hgc.stanford.edu

FEATURES

1..1381
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Comer Creek sticklebacks, WA USA"
/db_xref="taxon:69293"
/clone="CNB98-F11"
/sex="mixed male and female"
/tissue_type="brain"
/dev_stage="adult"
/lab_host="DH10B (T1 phage resistant)"
/clone_id="SHGC-CNB"
/note="Vector: Express 1; Total and poly A+ RNA was

isolated from the indicated stickleback tissue, and a cDNA library was constructed in the Express 1 plasmid vector by Open Biosystems. First strand cDNA synthesis was primed with an 54 bp linker primer containing an oligo(dT) sequence preceded by a synthetic NotI site (first strand primer: 5'-GACTAGTTCTAGATCGGAGCGGCCCC(T)25-3'). Following second strand synthesis, cDNAs were made blunt at the end corresponding to the original 5' prime end of mRNA, and cloned directionally into the NotI and EcoRV sites of Express 1. Note that the EcoRV site is typically destroyed in the blunt end cloning, leaving a junction of the form 'xxxxATC' (where is ATC is the second half of the EcoRV site, and xxx is derived from the cDNA sequence). A map of the Express 1 vector is available at: http://www.openbiosystems.com/cdna_library_construction_fa_q.php#8 The primary library was transformed and amplified in DH10B (T1 phage resistant) bacteria. Clones available from Open Biosystems: <http://www.openbiosystems.com/stickleback>

ORIGIN

Query Match 12.8%; Score 51; DB 9; Length 1381;
Best Local Similarity 56.4%; Pred. No. 0.041; Mismatches 71; Indels 0; Gaps 0;
Matches 93; Conservative 1;

123 TACCTTTTGTACTGTTGATATATATATATATACATACACACACACACACACA 182
91 TAT 150
183 YATATTTTAACTGGGACATTAATGTACACCACTATTCAAAGTAAACTACTATCTCA 242
151 CATATTTATATATATATATCTCAAAATACATAGTCTCAGCAGCAAGAAAACGTCACATTCC 210
243 GAGTAAGACATTAATCTGCTATGTCATTAATCTAGACATTAAG 287
211 TTGATTAACCAAGAAAGTAAATATCTTCAATTAATAAAAAAAG 255

RESULT 48

AKI39565/c 1722 bp mRNA linear HTC 21-SEP-2005
LOCUS AKI39565
DEFINITION Mus musculus 2 cells egg cDNA, RIKEN full-length enriched library,
clone:B0200040.8 product:unclassified, full insert sequence.
ACCESSION AKI39565
VERSION AKI39565.1 GI:74137777
KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636

AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Suganara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159

JOURNAL

PUBMED 11042159
REFERENCE 3
Shibata, K., Itoh, M., Mizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kitanai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsunoto, H., Sakeguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E., Matsumi, K.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.

TITLES

RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multiplexed cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL PUBMED REFERENCE AUTHORS	Genome Res. 10 (11), 1757-1771 (2000)
	11076861
JOURNAL PUBMED REFERENCE AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aikawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S., Aikawa, K., Aikawa, M., Nishii, K., Kiyosawa, H., Kondo, S., Yamanka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Platschmann, W., Gaasterland, T., Gissi, C., King, B., Kochwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaio, I., Pesole, G., Quackenbush, J., Schriml, L.M., Stabli, F., Suzuki, R., Tomita, M., Wagner, L., Mashio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gastlisch, S., Hill, D., Hotman, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Momberts, P., Nordone, P., Rings, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, J., Wynshaw-Boris, A., Yoshida, K., Haegawa, Y., Kawaji, H., Kontsuki, S. and Hayashizaki, Y.
	RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium
JOURNAL PUBMED REFERENCE AUTHORS	Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)
	11217851
JOURNAL PUBMED REFERENCE AUTHORS	Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaio, I., Ose, N., Saito, R., Suzuki, H., Yamanka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brusic, V., Chochla, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackeon, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierzki, R.M., King, B.L., Kongaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Nunata, K., Okido, T., Pavan, W.J., Petrea, G., Pesole, G., Petrowsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B., Ringwald, M., Sandelin, A., Schneider, C., Sempke, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayata, N., Hirozane-Kishikawa, T., Komno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Aikawa, T., Fukuda, S., Hara, A., Hashizume, W., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shibata, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
	PANTOM Consortium
JOURNAL PUBMED REFERENCE AUTHORS	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420 (6915), 563-573 (2002)
	12466851
JOURNAL PUBMED REFERENCE AUTHORS	Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodaira, R., Shimokawa, K., Bajic, V.B., Bremner, S.B., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aldridge, J., Allen, J.E., Ambesi-Impombato, A., Apweiler, R., Aulic, A., Avila, R., Bailey, T.L., Bamshad, M., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Christoffels, A., Clutterbuck, D.R., Crome, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagiolini, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N.,
	Hill, D., Humnicki, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelsio, J., Kitamura, H., Kitano, H., Kollas, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Laveau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liu, S., McWilliam, S., Madan Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matuzawa, S., Miki, H., Mignone, F., Miyake, S., Morris, K., Mortrud, T., Muller, N., Nakano, N., Nakachi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, P., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavoni, G., Pesole, G., Petrowsky, N., Piazza, S., Reed, J., Reid, J.P., Ring, B.Z., Ringwald, M., Rott, B., Ruan, Y., Salzberg, S.L., Sandelin, A., Schneider, C., Schonbach, C., Sekituchi, K., Sempke, C.A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sincalir, B., Sperling, S., Stupka, E., Sugita, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R., Wei, C.L., Yagi, K., Yamashita, H., Zdobych, E., Zhu, S., Zimmer, A., Zimmer, A., Bult, C., Grimmond, S.M., Teasdale, R.D., Liu, E.T., Brusic, V., Quackenbush, J., Wahlstedt, C., Mattick, J.S., Hume, D.A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori, Katayama, M., Suzuki, M., Aoki, J., Aikawa, T., Iida, J., Imamura, K., Kojima, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Komno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessey, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watanishi, A., Okamura-Oho, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.
	FANTOM Consortium
JOURNAL PUBMED REFERENCE AUTHORS	The transcriptional landscape of the mammalian genome Science 309 (5740), 1559-1563 (2005)
	16141072
JOURNAL PUBMED REFERENCE AUTHORS	Karayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yagi, C.C., Suzuki, M., Kawai, J., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Engstrom, P.G., Mizuno, Y., Faghihi, M.A., Sandelin, A., Wahlstedt, C., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Riken Genome Exploration Research Group
	Antisense transcription in the mammalian transcriptome Science 309 (5740), 1564-1566 (2005)
JOURNAL PUBMED REFERENCE AUTHORS	16141073
	8 (bases 1 to 1722)
JOURNAL PUBMED REFERENCE AUTHORS	Arkawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hori, P., Iida, J., Imamura, K., Imocani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Komno, H., Murata, M., Nakamura, M., Ninomiya, N., Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Watanishi, A., Muramatsu, M. and Hayashizaki, Y.
	Direct Submission
JOURNAL PUBMED REFERENCE AUTHORS	Submitted (30-MAR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
JOURNAL PUBMED REFERENCE AUTHORS	Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ Location/Qualifiers
	1. 1722 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="PANTOM:DB:B020004018" /db_xref="taxon:10090" /clone="B020004018" /tissue_type="egg"

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OW nucleic - nucleic search, using sw model

Run on: July 17, 2006, 17:09:12 ; Search time 151 Seconds
(without alignments)
4956.580 Million cell updates/sec

Title: SEQ1-33670G

Perfect score: 399.2

Sequence: 1 aagatgctctctctctata.....taactcagatcagacgagag 400

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

Issued Patents NA:*

1: /EMC_Celerra_SIDS3/ptodata/2/ina/1 COMB.seq:*

2: /EMC_Celerra_SIDS3/ptodata/2/ina/5 COMB.seq:*

3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A COMB.seq:*

4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B COMB.seq:*

5: /EMC_Celerra_SIDS3/ptodata/2/ina/7 COMB.seq:*

6: /EMC_Celerra_SIDS3/ptodata/2/ina/H COMB.seq:*

7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCBUS COMB.seq:*

8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP COMB.seq:*

9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE COMB.seq:*

10: /EMC_Celerra_SIDS3/ptodata/2/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	52.4	13.1	601	US-09-949-016-122917	Sequence 122917,
2	52.4	13.1	601	US-09-949-016-123011	Sequence 123011,
3	52.4	13.1	601	US-09-949-016-123105	Sequence 123105,
4	52.4	13.1	601	US-09-949-016-123199	Sequence 123199,
5	52.4	13.1	601	US-09-949-016-123293	Sequence 123293,
6	52.4	13.1	3455	US-10-009-332-31	Sequence 31, Appl
7	51.2	12.8	601	US-09-949-016-122919	Sequence 122919,
8	51.2	12.8	601	US-09-949-016-122920	Sequence 122920,
9	51.2	12.8	601	US-09-949-016-122922	Sequence 122922,
10	51.2	12.8	601	US-09-949-016-123013	Sequence 123013,
11	51.2	12.8	601	US-09-949-016-123014	Sequence 123014,
12	51.2	12.8	601	US-09-949-016-123016	Sequence 123016,
13	51.2	12.8	601	US-09-949-016-123107	Sequence 123107,
14	51.2	12.8	601	US-09-949-016-123108	Sequence 123108,
15	51.2	12.8	601	US-09-949-016-123201	Sequence 123201,
16	51.2	12.8	601	US-09-949-016-123202	Sequence 123202,
17	51.2	12.8	601	US-09-949-016-123204	Sequence 123204,
18	51.2	12.8	601	US-09-949-016-123295	Sequence 123295,
19	51.2	12.8	601	US-09-949-016-123296	Sequence 123296,
20	51.2	12.8	601	US-09-949-016-123298	Sequence 123298,
21	51.2	12.8	601	US-09-949-016-15212	Sequence 15212, A
22	51.2	12.8	64377	US-09-949-016-15213	Sequence 15213, A
23	51.2	12.8	64377	US-09-949-016-15213	Sequence 15213, A

24	51.2	12.8	64377	US-09-949-016-15214	Sequence 15214, A
25	51.2	12.8	64377	US-09-949-016-15215	Sequence 15215, A
26	51.2	12.8	64377	US-09-949-016-15216	Sequence 15216, A
27	51	12.8	3462	US-10-009-332-30	Sequence 30, Appl
28	51	12.8	3464	US-10-009-332-26	Sequence 26, Appl
29	51	12.8	3467	US-10-009-332-25	Sequence 25, Appl
30	51	12.8	3467	US-10-009-332-29	Sequence 29, Appl
31	51	12.8	3469	US-10-009-332-27	Sequence 27, Appl
32	51	12.8	3470	US-10-009-332-28	Sequence 28, Appl
33	51	12.8	3473	US-10-009-332-24	Sequence 24, Appl
34	50.8	12.7	601	US-09-949-016-122915	Sequence 122915,
35	50.8	12.7	601	US-09-949-016-122916	Sequence 122916,
36	50.8	12.7	601	US-09-949-016-123009	Sequence 123009,
37	50.8	12.7	601	US-09-949-016-123010	Sequence 123010,
38	50.8	12.7	601	US-09-949-016-123103	Sequence 123103,
39	50.8	12.7	601	US-09-949-016-123104	Sequence 123104,
40	50.8	12.7	601	US-09-949-016-123197	Sequence 123197,
41	50.8	12.7	601	US-09-949-016-123198	Sequence 123198,
42	50.8	12.7	601	US-09-949-016-123291	Sequence 123291,
43	50.8	12.7	601	US-09-949-016-123292	Sequence 123292,
44	50	12.5	601	US-09-949-016-122921	Sequence 122921,
45	50	12.5	601	US-09-949-016-123015	Sequence 123015,
46	50	12.5	601	US-09-949-016-123109	Sequence 123109,
47	50	12.5	601	US-09-949-016-123203	Sequence 123203,
48	50	12.5	601	US-09-949-016-123297	Sequence 123297,
49	49.8	12.5	2534	US-09-526-193A-23	Sequence 23, Appl
50	49.8	12.5	153866	US-09-949-016-16919	Sequence 16919, A
51	49.6	12.4	601	US-09-949-016-82380	Sequence 82380, A
52	49.6	12.4	601	US-09-949-016-122918	Sequence 122918,
53	49.6	12.4	601	US-09-949-016-123016	Sequence 123016,
54	49.6	12.4	601	US-09-949-016-123204	Sequence 123204,
55	49.6	12.4	601	US-09-949-016-123294	Sequence 123294,
56	49.6	12.4	601	US-09-949-016-187550	Sequence 187550,
57	49.6	12.4	312470	US-09-949-016-14043	Sequence 14043, A
58	49.6	12.4	336024	US-09-949-016-12373	Sequence 12373, A
59	48.6	12.2	601	US-09-949-016-87162	Sequence 87162, A
60	48.6	12.2	140840	US-09-949-016-14199	Sequence 14199, A
61	48.6	12.2	786431	US-09-751-389-3	Sequence 3, Appl
62	48.6	12.2	601	US-09-949-016-82379	Sequence 82379, A
63	48.4	12.1	601	US-09-949-016-187549	Sequence 187549,
64	48.4	12.1	66933	US-09-544-398B-11	Sequence 11, Appl
65	48.4	12.1	66933	US-09-544-398B-11	Sequence 11, Appl
66	48.4	12.1	66933	US-09-544-398B-11	Sequence 11, Appl
67	48.4	12.1	72049	US-09-544-398B-9	Sequence 9, Appl
68	48.4	12.1	72049	US-09-544-398B-9	Sequence 9, Appl
69	48.4	12.1	98962	US-09-949-016-14133	Sequence 14133, A
70	48.4	12.1	102884	US-09-949-016-17100	Sequence 17100, A
71	48.2	12.1	134890	US-09-949-016-15602	Sequence 15602, A
72	48.2	12.1	183202	US-09-949-016-13614	Sequence 13614, A
73	48	12.0	601	US-09-949-016-56477	Sequence 56477, A
74	48	12.0	601	US-09-949-016-56478	Sequence 56478, A
75	48	12.0	601	US-09-949-016-56479	Sequence 56479, A
76	48	12.0	601	US-09-949-016-113290	Sequence 113290,
77	48	12.0	601	US-09-949-016-159266	Sequence 159266,
78	48	12.0	148794	US-09-949-016-12751	Sequence 12751, A
79	48	12.0	285986	US-09-949-016-12287	Sequence 12287, A
80	48	12.0	288031	US-09-949-016-14864	Sequence 14864, A
81	48	12.0	635591	US-09-949-016-113808	Sequence 113808, A
82	48	12.0	635591	US-09-949-016-13388	Sequence 13388, A
83	47.8	12.0	601	US-09-949-016-82382	Sequence 82382, A
84	47.8	12.0	601	US-09-949-016-187552	Sequence 187552, A
85	47.8	12.0	93778	US-09-949-016-15096	Sequence 15096, A
86	47.8	12.0	209631	US-09-949-002-574	Sequence 574, App
87	47.8	12.0	209631	US-09-949-002-802	Sequence 802, App
88	47.6	11.9	601	US-09-949-016-174360	Sequence 174360, A
89	47.6	11.9	101349	US-09-949-016-174363	Sequence 174363, A
90	47.6	11.9	117937	US-09-949-016-12762	Sequence 12762, A
91	47.6	11.9	117937	US-09-949-016-15775	Sequence 15775, A
92	47.6	11.9	122626	US-09-949-016-17524	Sequence 17524, A
93	47.4	11.9	601	US-09-949-016-112707	Sequence 112707, A
94	47.4	11.9	601	US-09-949-016-112709	Sequence 112709, A
95	47.4	11.9	601	US-09-949-016-112791	Sequence 112791, A
96	47.4	11.9	601	US-09-949-016-112793	Sequence 112793, A

```

C 97 47.4 11.9 601 3 US-09-949-016-112875 Sequence 112875,
C 98 47.4 11.9 601 3 US-09-949-016-112877 Sequence 112877,
C 99 47.4 11.9 601 3 US-09-949-016-112851 Sequence 112951,
C 100 47.4 11.9 601 3 US-09-949-016-112953 Sequence 112953,
C 101 47.4 11.9 601 3 US-09-949-016-113037 Sequence 113037,
C 102 47.4 11.9 601 3 US-09-949-016-113039 Sequence 113039,
C 103 47.4 11.9 106380 3 US-09-949-016-17553 Sequence 17553, A
C 104 47.4 11.9 127280 3 US-09-949-016-14857 Sequence 14857, A
C 105 47.4 11.9 122266 3 US-09-949-016-14860 Sequence 14860, A
C 106 47.4 11.9 150933 3 US-09-949-016-14859 Sequence 14859, A
C 107 47.4 11.9 159963 3 US-09-949-016-14858 Sequence 14858, A
C 108 47.4 11.9 171130 3 US-09-949-016-14861 Sequence 14861, A
C 109 47.4 11.9 256287 3 US-09-949-016-14668 Sequence 14668, A
C 110 47.4 11.9 360470 3 US-09-949-016-13173 Sequence 13173, A
C 111 47.2 11.8 60417 3 US-09-949-016-13312 Sequence 13312, A
C 112 47.2 11.8 321022 3 US-09-949-016-11852 Sequence 11852, A
C 113 47.2 11.8 321022 3 US-09-949-016-14166 Sequence 14166, A
C 114 47 11.8 601 3 US-09-949-016-112708 Sequence 112708,
C 115 47 11.8 601 3 US-09-949-016-112792 Sequence 112792,
C 116 47 11.8 601 3 US-09-949-016-112876 Sequence 112876,
C 117 47 11.8 601 3 US-09-949-016-112852 Sequence 112952,
C 118 47 11.8 601 3 US-09-949-016-113038 Sequence 113038,
C 119 47 11.8 23445 3 US-09-949-016-12955 Sequence 12955, A
C 120 47 11.8 818128 3 US-09-949-016-14546 Sequence 14546, A
C 121 47 11.8 818128 3 US-09-949-016-14547 Sequence 14547, A
C 122 47 11.8 818128 3 US-09-949-016-14548 Sequence 14548, A
C 123 47 11.8 818128 3 US-09-949-016-14549 Sequence 14549, A
C 124 47 11.8 818128 3 US-09-949-016-14550 Sequence 14550, A
C 125 47 11.8 818128 3 US-09-949-016-14551 Sequence 14551, A
C 126 47 11.8 818128 3 US-09-949-016-14552 Sequence 14552, A
C 127 47 11.8 818128 3 US-09-949-016-14553 Sequence 14553, A
C 128 47 11.8 818128 3 US-09-949-016-14554 Sequence 14554, A
C 129 47 11.8 818128 3 US-09-949-016-14555 Sequence 14555, A
C 130 47 11.8 818128 3 US-09-949-016-14556 Sequence 14556, A
C 131 47 11.8 818128 3 US-09-949-016-14557 Sequence 14557, A
C 132 47 11.8 818128 3 US-09-949-016-14558 Sequence 14558, A
C 133 47 11.8 818128 3 US-09-949-016-14559 Sequence 14559, A
C 134 47 11.8 818128 3 US-09-949-016-14560 Sequence 14560, A
C 135 47 11.8 818128 3 US-09-949-016-14561 Sequence 14561, A
C 136 47 11.8 818128 3 US-09-949-016-14562 Sequence 14562, A
C 137 47 11.8 818128 3 US-09-949-016-14564 Sequence 14564, A
C 138 47 11.8 818128 3 US-09-949-016-14565 Sequence 14565, A
C 139 47 11.8 818128 3 US-09-949-016-14566 Sequence 14566, A
C 140 47 11.8 818128 3 US-09-949-016-14567 Sequence 14567, A
C 141 46.8 11.7 601 3 US-09-949-016-24727 Sequence 24727, A
C 142 46.8 11.7 601 3 US-09-949-016-82381 Sequence 82381, A
C 143 46.8 11.7 601 3 US-09-949-016-149790 Sequence 149790,
C 144 46.8 11.7 601 3 US-09-949-016-187551 Sequence 187551,
C 145 46.8 11.7 78810 3 US-09-949-016-16198 Sequence 16198, A
C 146 46.8 11.7 146401 3 US-09-949-016-16151 Sequence 16151, A
C 147 46.8 11.7 167708 3 US-09-949-016-16423 Sequence 16423, A
C 148 46.8 11.7 194790 3 US-09-949-016-15393 Sequence 15393, A
C 149 46.6 11.7 601 3 US-09-949-016-60589 Sequence 60589, A
C 150 46.6 11.7 601 3 US-09-949-016-60590 Sequence 60590, A

```

ALIGNMENTS

```

RESULT 1
US-09-949-016-122917/c
; Sequence 122917, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

```

```

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 122917
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-122917

```

```

Query Match 13.1%; Score 52.4; DB 3; Length 601;
Best Local Similarity 60.0%; Pred. No. 4.5e-06;
Matches 102; Conservative 2; Mismatches 63; Indels 3; Gaps 1;

```

```

QY 88 CCTGTGAAAGTATGATGACAGGCAAGTACCTCTTTTGTACTGTTATATAT 147
DB 392 CATCAGGGCAGAGAGAGATTGAAAGCTACCTCAGTTGGTGTATTCATATAT 333
QY 148 ATATATACATACACACACACACACACACACACACACATATTTTAACTCGGACATAATG 207
DB 332 ATATATACACACACACACACACACACACACACACATATATATATATACAC--ACACATATA 276
QY 208 TACACCACTATTTCATAAATAAACTACTATCTCAGAGTAAGACATPAA 257
DB 275 TATTCATATATGGAATATGATATATATATATAGAAATATATGTA 226

```

```

RESULT 2
US-09-949-016-123011/c
; Sequence 123011, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

```

```

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 123011
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-123011

```

```

Query Match 13.1%; Score 52.4; DB 3; Length 601;
Best Local Similarity 60.0%; Pred. No. 4.5e-06;
Matches 102; Conservative 2; Mismatches 63; Indels 3; Gaps 1;

```

```

QY 88 CCTGTGAAAGTATGATGACAGGCAAGTACCTCTTTTGTACTGTTATATAT 147
DB 392 CATCAGGGCAGAGAGAGATTGAAAGCTACCTCAGTTGGTGTATTCATATAT 333
QY 148 ATATATACATACACACACACACACACACACACACACATATTTTAACTCGGACATAATG 207
DB 332 ATATATACACACACACACACACACACACACACACATATATATATATACAC--ACACATATA 276
QY 208 TACACCACTATTTCATAAATAAACTACTATCTCAGAGTAAGACATPAA 257
DB 275 TATTCATATATGGAATATGATATATATATATAGAAATATATGTA 226

```

```

RESULT 3
US-09-949-016-123105/c
; Sequence 123105, Application US/09949016

```

Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 123105
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-123105

Query Match
Best Local Similarity 13.1%; Score 52.4; DB 3; Length 601;
60.0%; Pred. No. 4.5e-06;
Matches 102; Conservative 2; Mismatches 63; Indels 3; Gaps 1;

88 CCTCTGGAAGTAGTATGACAGGCAAGTGGCTTCTTTTGTACTGTTGATATAT 147
DB CATCAGGGCAAGAGACAGATTTGAAAGCTTCTCAGTTGGGTGTTATTTCTATATAT 333

148 ATATATACATACACACACACACACACACACACACACATATTTTAACTCGGGACATAATG 207
DB ATATATACACACACACACACACACACACACACACATATATATATATATACAC--ACACATATA 276

208 TACACCACTTATTCAGTAAGTAACTACTATCTCAGAGTAAGACATPA 257
DB TATTCATATATGAAATGTATATATATATATATATATAGAAATATATGA 226

RESULT 4
US-09-949-016-123199/c
Sequence 123199, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 123199
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-123199

Query Match
Best Local Similarity 13.1%; Score 52.4; DB 3; Length 601;
60.0%; Pred. No. 4.5e-06;
Matches 102; Conservative 2; Mismatches 63; Indels 3; Gaps 1;

88 CCTCTGGAAGTAGTATGACAGGCAAGTGGCTTCTTTTGTACTGTTGATATAT 147
DB CATCAGGGCAAGAGACAGATTTGAAAGCTTCTCAGTTGGGTGTTATTTCTATATAT 333

148 ATATATACATACACACACACACACACACACACACACATATTTTAACTCGGGACATAATG 207
DB ATATATACATACACACACACACACACACACACACACATATATATATATATACAC--ACACATATA 276

208 TACACCACTTATTCAGTAAGTAACTACTATCTCAGAGTAAGACATPA 257
DB TATTCATATATGAAATGTATATATATATATATATATAGAAATATATGA 226

DB 332 ATATATACATACACACACACACACACACACACACATATATATATATATACAC--ACACATATA 276

208 TACACCACTTATTCAGTAAGTAACTACTATCTCAGAGTAAGACATPA 257
DB TATTCATATATGAAATGTATATATATATATATATATAGAAATATATGA 226

RESULT 5
US-09-949-016-123293/c
Sequence 123293, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 123293
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-123293

Query Match
Best Local Similarity 13.1%; Score 52.4; DB 3; Length 601;
60.0%; Pred. No. 4.5e-06;
Matches 102; Conservative 2; Mismatches 63; Indels 3; Gaps 1;

88 CCTCTGGAAGTAGTATGACAGGCAAGTGGCTTCTTTTGTACTGTTGATATAT 147
DB CATCAGGGCAAGAGACAGATTTGAAAGCTTCTCAGTTGGGTGTTATTTCTATATAT 333

148 ATATATACATACACACACACACACACACACACACACATATTTTAACTCGGGACATAATG 207
DB ATATATACATACACACACACACACACACACACACACATATATATATATATACAC--ACACATATA 276

208 TACACCACTTATTCAGTAAGTAACTACTATCTCAGAGTAAGACATPA 257
DB TATTCATATATGAAATGTATATATATATATATATATAGAAATATATGA 226

RESULT 6
US-10-009-332-31/c
Sequence 31, Application US/10009332
Patent No. 6716613
GENERAL INFORMATION:
APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING AGGRECANASE ACTIVITY
FILE REFERENCE: 067541
CURRENT APPLICATION NUMBER: US/10/009,332
PRIOR FILING DATE: 2001-12-10
PRIOR APPLICATION NUMBER: JPA Hei 11-321740
PRIOR FILING DATE: 1999-11-11
PRIOR APPLICATION NUMBER: JPA 2000-144020
PRIOR FILING DATE: 2000-05-16
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.1
SEQ ID NO 31
LENGTH: 3455
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: promoter

Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123013
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-123013

Query Match 12.8%; Score 51.2; DB 3; Length 601;
Best Local Similarity 60.0%; Pred. No. 1e-05;
Matches 102; Conservative 1; Mismatches 64; Indels 3; Gaps 1;

88 CCTCTGAAAGTAGTATGACAGGCGAAGTGGAGGCTTTTGTACTGTTGATATAT 147
DB CATCAGGGCAAGAGACAGATTGTAAGCTCTCTCAGTTGGTGTATTTCTATATAT 345

148 ATATATACATACACACACACACACACACACACATATTTTAACTCGGGACATTAATG 207
DB ATATATACACACACACACACACACACACATATATATATATATACAC---ACACATATA 288

208 TACACCACTATTTCAGTAAGTAAACTACTATCTCAGAGTAAGACATATA 257
DB TATTCATATATGGAATATGTATATATATATATAGAAATATATATGA 238

RESULT 11
US-09-949-016-123014/c
; Sequence 123014, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123014
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-123014

Query Match 12.8%; Score 51.2; DB 3; Length 601;
Best Local Similarity 60.0%; Pred. No. 1e-05;
Matches 102; Conservative 1; Mismatches 64; Indels 3; Gaps 1;

88 CCTCTGAAAGTAGTATGACAGGCGAAGTGGAGGCTTTTGTACTGTTGATATAT 147
DB CATCAGGGCAAGAGACAGATTGTAAGCTCTCTCAGTTGGTGTATTTCTATATAT 346

148 ATATATACATACACACACACACACACACACATATTTTAACTCGGGACATTAATG 207

DB 345 ATATATACACACACACACACACACACACATATATATATATATACAC---ACACATATA 289

208 TACACCACTATTTCAGTAAGTAAACTACTATCTCAGAGTAAGACATATA 257
DB TATTCATATATGGAATATGTATATATATATATAGAAATATATATGA 239

RESULT 12
US-09-949-016-123016/c
; Sequence 123016, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123016
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-123016

Query Match 12.8%; Score 51.2; DB 3; Length 601;
Best Local Similarity 60.0%; Pred. No. 1e-05;
Matches 102; Conservative 1; Mismatches 64; Indels 3; Gaps 1;

88 CCTCTGAAAGTAGTATGACAGGCGAAGTGGAGGCTTTTGTACTGTTGATATAT 147
DB CATCAGGGCAAGAGACAGATTGTAAGCTCTCTCAGTTGGTGTATTTCTATATAT 387

148 ATATATACATACACACACACACACACACACATATTTTAACTCGGGACATTAATG 207
DB ATATATACACACACACACACACACACATATATATATATATATACAC---ACACATATA 330

208 TACACCACTATTTCAGTAAGTAAACTACTATCTCAGAGTAAGACATATA 257
DB TATTCATATATGGAATATGTATATATATATATAGAAATATATATGA 280

RESULT 13
US-09-949-016-123107/c
; Sequence 123107, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123107
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human

US-09-949-016-123107

Query Match 12.8%; Score 51.2; DB 3; Length 601;

Best Local Similarity 60.0%; Pred. No. 1e-05; Mismatches 64; Indels 3; Gaps 1;

Matches 102; Conservative 1; Mismatches 64; Indels 3; Gaps 1;

QY 88 CCTCTGGAAGTAGTATGCAAGGCAAGTGAAGGCTTCTTTTGTACTGTTGATATAT 147
DB 404 CATCAGGGCAAGAGAGAGATTGAAAGCTACCTCAGTTGGGTGTATTTCTATATAT 345
QY 148 ATATATATATACATACACACACACACACACACATATATTTTAAGTGGGACATAATG 207
DB 344 ATATATATACACACACACACACACACATATATATATATATATACAC--ACACATATA 288
QY 208 TACACCACTATTCAAGTAAACTATCTATCTAGAGATTAAGACATAA 257
DB 287 TATTCATATATGGAATATGATATATATATATATAGATATATAGAAATATATGA 238

RESULT 14

US-09-949-016-123108/c

Sequence 123108, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C0001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/231,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 123108

LENGTH: 601

TYPE: DNA

ORGANISM: Human

US-09-949-016-123108

Query Match 12.8%; Score 51.2; DB 3; Length 601;

Best Local Similarity 60.0%; Pred. No. 1e-05; Mismatches 64; Indels 3; Gaps 1;

Matches 102; Conservative 1; Mismatches 64; Indels 3; Gaps 1;

QY 88 CCTCTGGAAGTAGTATGCAAGGCAAGTGAAGGCTTCTTTTGTACTGTTGATATAT 147
DB 405 CATCAGGGCAAGAGAGAGATTGAAAGCTACCTCAGTTGGGTGTATTTCTATATAT 346
QY 148 ATATATATATACATACACACACACACACACATATATTTTAAGTGGGACATAATG 207
DB 345 ATATATATACACACACACACACACACATATATATATATATATACAC--ACACATATA 289
QY 208 TACACCACTATTCAAGTAAACTATCTATCTAGAGATTAAGACATAA 257
DB 287 TATTCATATATGGAATATGATATATATATATATAGATATATAGAAATATATGA 239

RESULT 15

US-09-949-016-123110/c

Sequence 123110, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C0001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 123110

LENGTH: 601

TYPE: DNA

ORGANISM: Human

US-09-949-016-123110

Query Match 12.8%; Score 51.2; DB 3; Length 601;

Best Local Similarity 60.0%; Pred. No. 1e-05; Mismatches 64; Indels 3; Gaps 1;

Matches 102; Conservative 1; Mismatches 64; Indels 3; Gaps 1;

QY 88 CCTCTGGAAGTAGTATGCAAGGCAAGTGAAGGCTTCTTTTGTACTGTTGATATAT 147
DB 446 CATCAGGGCAAGAGAGAGATTGAAAGCTACCTCAGTTGGGTGTATTTCTATATAT 387
QY 148 ATATATATATACATACACACACACACACACATATTTTAAGTGGGACATAATG 207
DB 386 ATATATATACACACACACACACACACATATATATATATATATACAC--ACACATATA 330
QY 208 TACACCACTATTCAAGTAAACTATCTATCTAGAGATTAAGACATAA 257
DB 329 TATTCATATATGGAATATGATATATATATATATAGATATATAGAAATATATGA 280

RESULT 16

US-09-949-016-123201/c

Sequence 123201, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C0001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 123201

LENGTH: 601

TYPE: DNA

ORGANISM: Human

US-09-949-016-123201

Query Match 12.8%; Score 51.2; DB 3; Length 601;

Best Local Similarity 60.0%; Pred. No. 1e-05; Mismatches 64; Indels 3; Gaps 1;

Matches 102; Conservative 1; Mismatches 64; Indels 3; Gaps 1;

QY 88 CCTCTGGAAGTAGTATGCAAGGCAAGTGAAGGCTTCTTTTGTACTGTTGATATAT 147
DB 404 CATCAGGGCAAGAGAGAGATTGAAAGCTACCTCAGTTGGGTGTATTTCTATATAT 345
QY 148 ATATATATATACATACACACACACACACACATATTTTAAGTGGGACATAATG 207
DB 344 ATATATATACACACACACACACACACATATATATATATATATATACAC--ACACATATA 288
QY 208 TACACCACTATTCAAGTAAACTATCTATCTAGAGATTAAGACATAA 257
DB 287 TATTCATATATGGAATATGATATATATATATATAGATATATAGAAATATATGA 238

RESULT 17

```
US-09-949-016-123202/c
: Sequence 123202, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: FILE REFERENCE: C1001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: PRIOR FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 123202
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Human
US-09-949-016-123202

Query Match      12.8%; Score 51.2; DB 3; Length 601;
Best Local Similarity 60.0%; Pred. No. 1e-05;
Matches 102; Conservative 1; Mismatches 64; Indels 3; Gaps 1;

QY 88 CCTCTGGAAGTAGTATGACGAGGCAAGTGAAGCCTCTTTTGTACTGTTGATATAT 147
DB 405 CATCAGGGCAAGAGACAGATTGTAAGCTACCTCTCAGTTGGGTGTATTCTATATAT 346
QY 148 ATATATATACATACACACACACACACACACACATATATTTTAACTCGGACATTAATG 207
DB 345 ATATATATACACACACACACACACACACATATATATATATATATACAC--ACACATATA 289
QY 208 TACACCACTATTTCAGTAAGTAAACTACTATCTCAGAAAGTAAAGCATTA 257
DB 288 TATTCATATATGGAATATGTATATATATATATAGAAATATATATGA 239

RESULT 18
US-09-949-016-123204/c
: Sequence 123204, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: FILE REFERENCE: C1001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: PRIOR FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 123204
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Human
US-09-949-016-123204

Query Match      12.8%; Score 51.2; DB 3; Length 601;
Best Local Similarity 60.0%; Pred. No. 1e-05;
Matches 102; Conservative 1; Mismatches 64; Indels 3; Gaps 1;
```

```
QY 148 ATATATATACATACACACACACACACACACATATATTTTAACTCGGACATTAATG 207
DB 386 ATATATATACACACACACACACACACATATATATATATATATATACAC--ACACATATA 330
QY 208 TACACCACTATTTCAGTAAGTAAACTACTATCTCAGAAAGTAAAGCATTA 257
DB 329 TATTCATATATGGAATATGTATATATATATATAGAAATATATATATGA 280

RESULT 19
US-09-949-016-123295/c
: Sequence 123295, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: FILE REFERENCE: C1001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: PRIOR FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 123295
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Human
US-09-949-016-123295

Query Match      12.8%; Score 51.2; DB 3; Length 601;
Best Local Similarity 60.0%; Pred. No. 1e-05;
Matches 102; Conservative 1; Mismatches 64; Indels 3; Gaps 1;

QY 88 CCTCTGGAAGTAGTATGACGAGGCAAGTGAAGCCTCTTTTGTACTGTTGATATAT 147
DB 404 CATCAGGGCAAGAGACAGATTGTAAGCTACCTCTCAGTTGGGTGTATTCTATATAT 345
QY 148 ATATATATACATACACACACACACACACACATATATTTTAACTCGGACATTAATG 207
DB 344 ATATATATACACACACACACACACACATATATATATATATATATACAC--ACACATATA 288
QY 208 TACACCACTATTTCAGTAAGTAAACTACTATCTCAGAAAGTAAAGCATTA 257
DB 287 TATTCATATATGGAATATGTATATATATATATAGAAATATATATATGA 238

RESULT 20
US-09-949-016-123296/c
: Sequence 123296, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: FILE REFERENCE: C1001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: PRIOR FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 123296
: LENGTH: 601
```

TYPE: DNA
ORGANISM: Human
US-09-949-016-123296

Query Match 12.8%; Score 51.2; DB 3; Length 601;
Best Local Similarity 60.0%; Pred. No. 1e-05;
Matches 102; Conservative 1; Mismatches 64; Indels 3; Gaps 1;

88 CCTCTGGAAGTAGTATGACAGGAGGAGGCTCTCTTTTGTACTGTTGATATAT 147
405 CATCAGGGCAAGAGACAGATTTTAAAGCTACCTCTCAGTTGGGTGTTATCTATATAT 346
148 ATATATACATACACACACACACACACACACACACACATATTTTAACTCGGACATAATG 207
345 ATATATACACACACACACACACACACACACATATATATATATATACAC---ACACATATA 289
208 TACACCACTTATTCAGTAAGTAAACTCTATCTCAGAGTAAGACATTA 257
288 TATTCATATATGGAATATGTATATATATATATAGAAATATATAGA 239

RESULT 21

US-09-949-016-123298/c
Sequence 123298, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: C1001307

CURRENT FILING DATE: 2000-04-14

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 123298

LENGTH: 601

TYPE: DNA

ORGANISM: Human

US-09-949-016-123298

Query Match 12.8%; Score 51.2; DB 3; Length 601;
Best Local Similarity 60.0%; Pred. No. 1e-05;
Matches 102; Conservative 1; Mismatches 64; Indels 3; Gaps 1;

88 CCTCTGGAAGTAGTATGACAGGAGGAGGCTCTCTTTTGTACTGTTGATATAT 147
446 CATCAGGGCAAGAGACAGATTTTAAAGCTACCTCTCAGTTGGGTGTTATCTATATAT 387
148 ATATATACATACACACACACACACACACACACACACATATTTTAACTCGGACATAATG 207
386 ATATATACACACACACACACACACACACACATATATATATATATATACAC---ACACATATA 330
208 TACACCACTTATTCAGTAAGTAAACTCTATCTCAGAGTAAGACATTA 257
329 TATTCATATATGGAATATGTATATATATATATAGAAATATATAGA 280

RESULT 22

US-09-949-016-15212

Sequence 15212, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: C1001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15212

LENGTH: 64377
TYPE: DNA
ORGANISM: Human
US-09-949-016-15212

Query Match 12.8%; Score 51.2; DB 3; Length 64377;
Best Local Similarity 60.0%; Pred. No. 6.7e-05;
Matches 102; Conservative 1; Mismatches 64; Indels 3; Gaps 1;

88 CCTCTGGAAGTAGTATGACAGGAGGAGGCTCTCTTTTGTACTGTTGATATAT 147
24190 CATCAGGGCAAGAGACAGATTTTAAAGCTACCTCTCAGTTGGGTGTTATCTATATAT 24249
148 ATATATACATACACACACACACACACACACACACACATATTTTAACTCGGACATAATG 207
24250 ATATATACACACACACACACACACACACACATATATATATATATACAC---ACACATATA 24306
208 TACACCACTTATTCAGTAAGTAAACTCTATCTCAGAGTAAGACATTA 257
24307 TATTCATATATGGAATATGTATATATATATATAGAAATATATAGA 24356

RESULT 23

US-09-949-016-15213

Sequence 15213, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: C1001307

CURRENT FILING DATE: 2000-04-14

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 15213

LENGTH: 64377

TYPE: DNA

ORGANISM: Human

US-09-949-016-15213

Query Match 12.8%; Score 51.2; DB 3; Length 64377;
Best Local Similarity 60.0%; Pred. No. 6.7e-05;
Matches 102; Conservative 1; Mismatches 64; Indels 3; Gaps 1;

88 CCTCTGGAAGTAGTATGACAGGAGGAGGCTCTCTTTTGTACTGTTGATATAT 147
24190 CATCAGGGCAAGAGACAGATTTTAAAGCTACCTCTCAGTTGGGTGTTATCTATATAT 24249
148 ATATATACATACACACACACACACACACACACACACATATTTTAACTCGGACATAATG 207
24250 ATATATACACACACACACACACACACACACATATATATATATATATACAC---ACACATATA 24306
208 TACACCACTTATTCAGTAAGTAAACTCTATCTCAGAGTAAGACATTA 257
24307 TATTCATATATGGAATATGTATATATATATATAGAAATATATAGA 24356

RESULT 24
US-09-949-016-15214
Sequence 15214, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15214
LENGTH: 64377
ORGANISM: Human
US-09-949-016-15214

Query Match
Best Local Similarity 12.8%; Score 51.2; DB 3; Length 64377;
Best Local Similarity 60.0%; Pred. No. 6.7e-05;
Matches 102; Conservative 1; Mismatches 64; Indels 3; Gaps 1;

Db
24190 CATCGGGCAGAGGACGAGTTTGAAGCTACCTTCTGCTTGGTGTATTTATATAT 24249
148 ATATATACATACACACACACACACACACATATATATATATATATATATATATAT 207
24250 ATATATACACACACACACACACACACATATATATATATATATATATATATATAT 24306
208 TACACCACTTTTCAAGTAAACTACTATCTCAGAAAGTAAAGCATATA 257
24307 TATTCATATATGGAATATGTATATATATATATATATATATATATATATATATAT 24356

RESULT 25
US-09-949-016-15215
Sequence 15215, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15215
LENGTH: 64377
ORGANISM: Human
US-09-949-016-15215

Query Match
Best Local Similarity 12.8%; Score 51.2; DB 3; Length 64377;
Best Local Similarity 60.0%; Pred. No. 6.7e-05;
Matches 102; Conservative 1; Mismatches 64; Indels 3; Gaps 1;

Db
24190 CATCGGGCAGAGGACGAGTTTGAAGCTACCTTCTGCTTGGTGTATTTATATAT 24249
148 ATATATACATACACACACACACACACACATATATATATATATATATATATATAT 207
24250 ATATATACACACACACACACACACACATATATATATATATATATATATATATAT 24306
208 TACACCACTTTTCAAGTAAACTACTATCTCAGAAAGTAAAGCATATA 257
24307 TATTCATATATGGAATATGTATATATATATATATATATATATATATATATATAT 24356

Db
24190 CATCGGGCAGAGGACGAGTTTGAAGCTACCTTCTGCTTGGTGTATTTATATAT 24249
148 ATATATACATACACACACACACACACACATATATATATATATATATATATATAT 207
24250 ATATATACACACACACACACACACACATATATATATATATATATATATATATAT 24306
208 TACACCACTTTTCAAGTAAACTACTATCTCAGAAAGTAAAGCATATA 257
24307 TATTCATATATGGAATATGTATATATATATATATATATATATATATATATATAT 24356

RESULT 26
US-09-949-016-15216
Sequence 15216, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15216
LENGTH: 64377
ORGANISM: Human
US-09-949-016-15216

Query Match
Best Local Similarity 12.8%; Score 51.2; DB 3; Length 64377;
Best Local Similarity 60.0%; Pred. No. 6.7e-05;
Matches 102; Conservative 1; Mismatches 64; Indels 3; Gaps 1;

Db
24190 CATCGGGCAGAGGACGAGTTTGAAGCTACCTTCTGCTTGGTGTATTTATATAT 24249
148 ATATATACATACACACACACACACACACATATATATATATATATATATATATAT 207
24250 ATATATACACACACACACACACACACATATATATATATATATATATATATATAT 24306
208 TACACCACTTTTCAAGTAAACTACTATCTCAGAAAGTAAAGCATATA 257
24307 TATTCATATATGGAATATGTATATATATATATATATATATATATATATATATAT 24356

RESULT 27
US-10-009-332-30/C
Sequence 30, Application US/10009332
Patent No. 6716613
GENERAL INFORMATION:
APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
TITLE OF INVENTION: NOVEL METALLOPROTEIN HAVING AGGRECANASE ACTIVITY
FILE REFERENCE: Q67541
CURRENT FILING DATE: 2001-12-10
PRIOR APPLICATION NUMBER: JPA Hei 11-321740
PRIOR FILING DATE: 1999-11-11
PRIOR APPLICATION NUMBER: JPA 2000-144020
PRIOR FILING DATE: 2000-05-16
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.1
SEQ ID NO 30
LENGTH: 3462

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: promoter
LOCATION: (1)..(3462)
OTHER INFORMATION:
US-10-009-332-30

Query Match 12.8%; Score 51; DB 3; Length 3462;
Best Local Similarity 82.6%; Pred. No. 2.4e-05;
Matches 57; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 121 CCTACCTTTTGTACTGTTGATATATATATATATACATACACACACACACACACA 180
DB 2414 CATACATATATATACATACATATATATATATATATATACATACACACACACACACA 2353

QY 181 CAYATATTT 189
DB 2352 CATATATAT 2344

RESULT 28
US-10-009-332-26/c
Sequence 26, Application US/10009332
Patent No. 6716613
GENERAL INFORMATION:
APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING AGGREGANASE ACTIVITY
FILE REFERENCE: Q67541
CURRENT APPLICATION NUMBER: US/10/009,332
CURRENT FILING DATE: 2001-12-10
PRIOR APPLICATION NUMBER: JPA Hei 11-321740
PRIOR FILING DATE: 1999-11-11
PRIOR APPLICATION NUMBER: JPA 2000-144020
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.1
SEQ ID NO 26
LENGTH: 3464
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: promoter
LOCATION: (1)..(3464)
OTHER INFORMATION:
US-10-009-332-26

Query Match 12.8%; Score 51; DB 3; Length 3464;
Best Local Similarity 82.6%; Pred. No. 2.4e-05;
Matches 57; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 121 CCTACCTTTTGTACTGTTGATATATATATATATACATACACACACACACACA 180
DB 2414 CATACATATATATACATACATATATATATATATATATACATACACACACACACACA 2355

QY 181 CAYATATTT 189
DB 2354 CATATATAT 2346

RESULT 29
US-10-009-332-25/c
Sequence 25, Application US/10009332
Patent No. 6716613
GENERAL INFORMATION:
APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING AGGREGANASE ACTIVITY
FILE REFERENCE: Q67541
CURRENT APPLICATION NUMBER: US/10/009,332
CURRENT FILING DATE: 2001-12-10
PRIOR APPLICATION NUMBER: JPA Hei 11-321740

PRIOR FILING DATE: 1999-11-11
PRIOR APPLICATION NUMBER: JPA 2000-144020
PRIOR FILING DATE: 2000-05-16
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.1
SEQ ID NO 25
LENGTH: 3467
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: promoter
LOCATION: (1)..(3467)
OTHER INFORMATION:
US-10-009-332-25

Query Match 12.8%; Score 51; DB 3; Length 3467;
Best Local Similarity 82.6%; Pred. No. 2.4e-05;
Matches 57; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 121 CCTACCTTTTGTACTGTTGATATATATATATATACATACACACACACACACA 180
DB 2414 CATACATATATATACATACATATATATATATATATATACATACACACACACACACA 2355

QY 181 CAYATATTT 189
DB 2354 CATATATAT 2346

RESULT 30
US-10-009-332-29/c
Sequence 29, Application US/10009332
Patent No. 6716613
GENERAL INFORMATION:
APPLICANT: Kazusa DNA Research Institute
TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING AGGREGANASE ACTIVITY
FILE REFERENCE: Q67541
CURRENT APPLICATION NUMBER: US/10/009,332
CURRENT FILING DATE: 2001-12-10
PRIOR APPLICATION NUMBER: JPA Hei 11-321740
PRIOR FILING DATE: 1999-11-11
PRIOR APPLICATION NUMBER: JPA 2000-144020
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.1
SEQ ID NO 29
LENGTH: 3467
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: promoter
LOCATION: (1)..(3467)
OTHER INFORMATION:
US-10-009-332-29

Query Match 12.8%; Score 51; DB 3; Length 3467;
Best Local Similarity 82.6%; Pred. No. 2.4e-05;
Matches 57; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 121 CCTACCTTTTGTACTGTTGATATATATATATATACATACACACACACACACA 180
DB 2418 CATACATATATATACATACATATATATATATATATATACATACACACACACACACA 2359

QY 181 CAYATATTT 189
DB 2358 CACATATAT 2350

RESULT 31
US-10-009-332-27/c
Sequence 27, Application US/10009332
Patent No. 6716613
GENERAL INFORMATION:

Db 368 CATCAGGCAAGAGGACAGATTTGAAAGCTACTCTCAGTTGGGTGTATTTATATAT 309
| | | | |
Qy 148 ATATATATATACATACACACACACACACACACATATTTTAACTCGGACATTAATG 207
| | | | |
Db 308 ATATATATATACACACACACACACACACATATATATATATATATACAC---ACACATATA 252
| | | | |
Qy 208 TACACCACTTATTCAAAGTAAATACTACTATCTCAGAGTAAGACATTA 257
| | | | |
Db 251 TATTCATATATGGAATATGTATATATATATATATAGAAATATAGATATATGA 202
| | | | |

RESULT 35
US-09-949-016-122916/c
; Sequence 122916, Application US/09949016
; Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: C1001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 122916

LENGTH: 601

TYPE: DNA

ORGANISM: Human

US-09-949-016-122916

Query Match 12.7%; Score 50.8; DB 3; Length 601;
Best Local Similarity 59.4%; Pred. No. 1.4e-05;

Matches 101; Conservative 2; Mismatches 64; Indels 3; Gaps 1;

Qy 88 CCTCTGAAAGTAGTATGACAGGCAAGTGAAGCCTCTTCTTGTACTGTTGATATAT 147
| | | | |
Db 390 CATCAGGCAAGAGGACAGATTTGAAAGCTACTCTCAGTTGGGTGTATTTATATAT 331
| | | | |
Qy 148 ATATATATATACATACACACACACACACACACATATTTTAACTCGGACATTAATG 207
| | | | |
Db 330 ATATATATATACACACACACACACACACATATATATATATATATACAC---ACACATATA 274
| | | | |
Qy 208 TACACCACTTATTCAAAGTAAATACTACTATCTCAGAGTAAGACATTA 257
| | | | |
Db 273 TATTCATATATGGAATATGTATATATATATATATAGAAATATAGATATATGA 224
| | | | |

RESULT 36
US-09-949-016-123009/c
; Sequence 123009, Application US/09949016
; Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: C1001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 123009

LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-123009

Query Match 12.7%; Score 50.8; DB 3; Length 601;
Best Local Similarity 59.4%; Pred. No. 1.4e-05;

Matches 101; Conservative 2; Mismatches 64; Indels 3; Gaps 1;

Qy 88 CCTCTGAAAGTAGTATGACAGGCAAGTGAAGCCTCTTCTTGTACTGTTGATATAT 147
| | | | |
Db 368 CATCAGGCAAGAGGACAGATTTGAAAGCTACTCTCAGTTGGGTGTATTTATATAT 309
| | | | |
Qy 148 ATATATATATACATACACACACACACACACACATATTTTAACTCGGACATTAATG 207
| | | | |
Db 308 ATATATATATACACACACACACACACACATATATATATATATATACAC---ACACATATA 252
| | | | |
Qy 208 TACACCACTTATTCAAAGTAAATACTACTATCTCAGAGTAAGACATTA 257
| | | | |
Db 251 TATTCATATATGGAATATGTATATATATATATAGAAATATAGATATATGA 202
| | | | |

RESULT 37
US-09-949-016-123010/c
; Sequence 123010, Application US/09949016
; Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: C1001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 123010

LENGTH: 601

TYPE: DNA

ORGANISM: Human

US-09-949-016-123010

Query Match 12.7%; Score 50.8; DB 3; Length 601;
Best Local Similarity 59.4%; Pred. No. 1.4e-05;

Matches 101; Conservative 2; Mismatches 64; Indels 3; Gaps 1;

Qy 88 CCTCTGAAAGTAGTATGACAGGCAAGTGAAGCCTCTTCTTGTACTGTTGATATAT 147
| | | | |
Db 390 CATCAGGCAAGAGGACAGATTTGAAAGCTACTCTCAGTTGGGTGTATTTATATAT 331
| | | | |
Qy 148 ATATATATATACATACACACACACACACACACATATTTTAACTCGGACATTAATG 207
| | | | |
Db 330 ATATATATATACACACACACACACACACATATATATATATATATACAC---ACACATATA 274
| | | | |
Qy 208 TACACCACTTATTCAAAGTAAATACTACTATCTCAGAGTAAGACATTA 257
| | | | |
Db 273 TATTCATATATGGAATATGTATATATATATATAGAAATATAGATATATGA 224
| | | | |

RESULT 38
US-09-949-016-123103/c
; Sequence 123103, Application US/09949016
; Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: C1001307


```

CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0.
SEQ ID NO 123103
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-123103

```

Query Match	12.7%	Score 50.8;	DB 3;	Length 601;
Best Local Similarity	59.4%;	Pred. No. 1.4e-05;		
Matches 101; Conservative	2;	Mismatches 64;	Indels 3;	Gaps 1

QY	88	CCCTCTGAAAGTGTGATTCAGGGCAGAGAGCCCTACCTTTTGTACTGTTTAAATAT	147
Db	368	CATCAGGGCAGAGGACAGATTGTAAGCTCCTTCACGTTTGGGTGTAATTCATATAT	309
QY	148	ATATATATACATACACACACACACACACATATTTTAACTCGGGACATATAATG	207
Db	308	ATATATATATACACACACACACACACACATATATATATATATATAC--ACACATATA	252
QY	208	TACACCACTTATTCAAAGTAAAACTACATCTCAGAGTAAAGCATTA	257
Db	251	TATTCATATATGGAATATGTATATATATATATGAAATTAATGAAATATATGA	202

RESULT 39
US-09-949-016-123104/c
; Sequence 123104, Application US/09949016

Query Match	12.7%	Score 50.8;	DB 3;	Length 601;
Best Local Similarity	59.4%	Pred. No. 1.4e-05;		
Matches 101; Conservative	2;	Mismatches 64;	Indels 3;	Gaps 1.

QY CCTCTGGAAGTAGGATGACAGGCGAATGAGGCTACTCTTTTGTACTGTTGATATAT 147
 Db 390 CACTGAGGGGCAAGGGACACATTTTGAACCTACTCTTCAGTTGGGATGTATTCTATATAT 3311
 QY 148 ATATATATACATACACACACACACACACACACATATTTTAATCTGGGACATAATG 207
 Db 330 ATATATATACACACACACACACACACACATATATATATATATATATACACACATATA 274
 QY 208 TACACGACCTATTCAAAGTAAAACTACATCTCTCGAAGTAAACACATAA 257
 Db 273 TATTCTATATATGAAATATGTATATATATATATATATATATATATAGAAATATATGA 224

RESULT 40
US-09-949-016-123197/c

Query Match	12.7%	Score 50.8	DB 3	Length 601
Best Local Similarity	59.4%	Pred. NO. 1.4e-05		
Matches 101; Conservative	2	Mismatches 64	Indels 3	Gaps 1

QY	88	CCCTCGAAGATAGGATATGACAGGGGAGATGAGCCCTACCTTTTGTGACGTTGATAT	147
Db	368	CATCAGGGCAAGAGGACAGATTTGAAACCTCACTCAAGTTTGGGTATTTCTATATAT	309
QY	148	ATATATCATATCACACACACACACACACACACATATTTTAACTGGGACATTAAG	207
Db	308	ATATATATACACACACACACACACACATATATATATATATAC---ACACATATA	252
QY	208	TACACGACCTATTCAAAGTAAAACTACTATCTCGAAGTAAAGACATTA	257
Db	251	TATTCCTATATATGGAATATGTATATATATATATAGAAATATAGATATATATGA	202

```

RESULT 41
US-09-949-016-123198/c
Sequence 123198 Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C0001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ. ID NOS: 207012
SOFTWARE: Rastcseq for Windows Version 4.0

```

Query Match	12.7%	Score 50.8;	DB 3;	Length 601;
Best Local Similarity	59.4%	Pred. No. 1.4e-05;		
Matches 101; Conservative	2;	Mismatches 64;	Indels 3;	Gaps 1;

Query 88 CCTCTGGAAGTAGTATGACAGGCAAGTACCTTTTGTACTGTTGATATAT 147
; SEQ ID NO 123292
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-123292
Db 390 CATCAGGGCAAGAGACAGATTGTAAGCTACCTCAGTTGGGTGTATTTATATAT 331
Qy 148 ATATATATCATACACACACACACACACACACACATATTTTAACTCGGACATAATG 207
Db 330 ATATATATACACACACACACACACACACACATATATATATATATATACAC--ACACATATA 274
Qy 208 TACACCACTATTTCAAAGTAAACTACTATCTCAGAGTAAGACATTA 257
Db 273 TATTCATATATGGAATATGTATATATATATATATAGAAATATATGA 224

RESULT 42

US-09-949-016-123291/c
; Sequence 123291, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 123291
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-123291

Query Match 12.7%; Score 50.8; DB 3; Length 601;
Best Local Similarity 59.4%; Pred. No. 1.4e-05;
Matches 101; Conservative 2; Mismatches 64; Indels 3; Gaps 1;

Qy 88 CCTCTGGAAGTAGTATGACAGGCAAGTACCTTTTGTACTGTTGATATAT 147
Db 368 CATCAGGGCAAGAGACAGATTGTAAGCTACCTCAGTTGGGTGTATTTATATAT 309
Qy 148 ATATATATCATACACACACACACACACACACACATATTTTAACTCGGACATAATG 207
Db 308 ATATATATACACACACACACACACACACACATATATATATATATATACAC--ACACATATA 252
Qy 208 TACACCACTATTTCAAAGTAAACTACTATCTCAGAGTAAGACATTA 257
Db 251 TATTCATATATGGAATATGTATATATATATATATAGAAATATATGA 202

RESULT 43

US-09-949-016-123292/c
; Sequence 123292, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 123292
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-123292

Query Match 12.7%; Score 50.8; DB 3; Length 601;
Best Local Similarity 59.4%; Pred. No. 1.4e-05;
Matches 101; Conservative 2; Mismatches 64; Indels 3; Gaps 1;

Qy 88 CCTCTGGAAGTAGTATGACAGGCAAGTACCTTTTGTACTGTTGATATAT 147
Db 390 CATCAGGGCAAGAGACAGATTGTAAGCTACCTCAGTTGGGTGTATTTATATAT 331
Qy 148 ATATATATCATACACACACACACACACACACACATATTTTAACTCGGACATAATG 207
Db 330 ATATATATACACACACACACACACACACACATATATATATATATATACAC--ACACATATA 274
Qy 208 TACACCACTATTTCAAAGTAAACTACTATCTCAGAGTAAGACATTA 257
Db 273 TATTCATATATGGAATATGTATATATATATATATAGAAATATATGA 224

RESULT 44

US-09-949-016-122921/c
; Sequence 122921, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 122921
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-122921

Query Match 12.5%; Score 50; DB 3; Length 601;
Best Local Similarity 65.7%; Pred. No. 2.4e-05;
Matches 71; Conservative 1; Mismatches 36; Indels 0; Gaps 0;

Qy 88 CCTCTGGAAGTAGTATGACAGGCAAGTACCTTTTGTACTGTTGATATAT 147
Db 413 CATCAGGGCAAGAGACAGATTGTAAGCTACCTCAGTTGGGTGTATTTATATAT 354
Qy 148 ATATATATCATACACACACACACACACACACACATATTTTAACTCGGACATAATG 195
Db 353 ATATATATACACACACACACACACACACACATATATATATATATATACAC 306

RESULT 45

US-09-949-016-123015/c
; Sequence 123015, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 123015
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-123015

Query Match 12.5%; Score 50; DB 3; Length 601;
Best Local Similarity 65.7%; Pred. No. 2.4e-05;
Matches 71; Conservative 1; Mismatches 36; Indels 0; Gaps 0;

QY 88 CCTCTGGAAGTAGTATGACGGCAAGTGAAGCCCTCTTTTGTACTGTTGATATAT 147
DB 413 CATCAGGGCAGAGGACAGATTGAAAGCTACCTCTGAGTTGGGTGTTATTTCTATATAT 354
QY 148 ATATATACATACACACACACACACACACACACACATATATTTTAATC 195
DB 353 ATATATACACACACACACACACACACACACATATATATATATATACAC 306

RESULT 46
US-09-949-016-123109/c
Sequence 123109, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 123109
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-123109

Query Match 12.5%; Score 50; DB 3; Length 601;
Best Local Similarity 65.7%; Pred. No. 2.4e-05;
Matches 71; Conservative 1; Mismatches 36; Indels 0; Gaps 0;

QY 88 CCTCTGGAAGTAGTATGACGGCAAGTGAAGCCCTCTTTTGTACTGTTGATATAT 147
DB 413 CATCAGGGCAGAGGACAGATTGAAAGCTACCTCTGAGTTGGGTGTTATTTCTATATAT 354
QY 148 ATATATACATACACACACACACACACACACACACATATATTTTAATC 195
DB 353 ATATATACACACACACACACACACACACACATATATATATATATACAC 306

RESULT 47
US-09-949-016-123203/c
Sequence 123203, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 123203
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-123203

FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 123203
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-123203

Query Match 12.5%; Score 50; DB 3; Length 601;
Best Local Similarity 65.7%; Pred. No. 2.4e-05;
Matches 71; Conservative 1; Mismatches 36; Indels 0; Gaps 0;

QY 88 CCTCTGGAAGTAGTATGACGGCAAGTGAAGCCCTCTTTTGTACTGTTGATATAT 147
DB 413 CATCAGGGCAGAGGACAGATTGAAAGCTACCTCTGAGTTGGGTGTTATTTCTATATAT 354
QY 148 ATATATACATACACACACACACACACACACACACATATATTTTAATC 195
DB 353 ATATATACACACACACACACACACACACACATATATATATATATACAC 306

RESULT 48
US-09-949-016-123297/c
Sequence 123297, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 123297
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-123297

Query Match 12.5%; Score 50; DB 3; Length 601;
Best Local Similarity 65.7%; Pred. No. 2.4e-05;
Matches 71; Conservative 1; Mismatches 36; Indels 0; Gaps 0;

QY 88 CCTCTGGAAGTAGTATGACGGCAAGTGAAGCCCTCTTTTGTACTGTTGATATAT 147
DB 413 CATCAGGGCAGAGGACAGATTGAAAGCTACCTCTGAGTTGGGTGTTATTTCTATATAT 354
QY 148 ATATATACATACACACACACACACACACACACACATATATTTTAATC 195
DB 353 ATATATACACACACACACACACACACACACATATATATATATATACAC 306

RESULT 49
US-09-526-193A-23/c
Sequence 23, Application US/09526193A
Patent No. 6617122
GENERAL INFORMATION:

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

On nucleic - nucleic search, using sw model

Run on: July 17, 2006, 16:04:09 ; Search time 1439 Seconds

(without alignments)
3415.604 Million cell updates/sec

Title: SEQ1-33670G

Perfect score: 399.2
Sequence: 1 aagatgctctctctctata.....taactcagatcagaagag 400

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 150 summaries

Database : Published Applications NA Main:

```
1: /EMC_Celerra_SIDS3/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC_Celerra_SIDS3/prodata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC_Celerra_SIDS3/prodata/2/pubpna/US09_PUBCOMB.seq:*
4: /EMC_Celerra_SIDS3/prodata/2/pubpna/US09C_PUBCOMB.seq:*
5: /EMC_Celerra_SIDS3/prodata/2/pubpna/US10_PUBCOMB.seq:*
6: /EMC_Celerra_SIDS3/prodata/2/pubpna/US10A_PUBCOMB.seq:*
7: /EMC_Celerra_SIDS3/prodata/2/pubpna/US10B_PUBCOMB.seq:*
8: /EMC_Celerra_SIDS3/prodata/2/pubpna/US10C_PUBCOMB.seq:*
9: /EMC_Celerra_SIDS3/prodata/2/pubpna/US10D_PUBCOMB.seq:*
10: /EMC_Celerra_SIDS3/prodata/2/pubpna/US10E_PUBCOMB.seq:*
11: /EMC_Celerra_SIDS3/prodata/2/pubpna/US10F_PUBCOMB.seq:*
12: /EMC_Celerra_SIDS3/prodata/2/pubpna/US10G_PUBCOMB.seq:*
13: /EMC_Celerra_SIDS3/prodata/2/pubpna/US11_PUBCOMB.seq:*
14: /EMC_Celerra_SIDS3/prodata/2/pubpna/US11A_PUBCOMB.seq:*
15: /EMC_Celerra_SIDS3/prodata/2/pubpna/US11C_PUBCOMB.seq:*
16: /EMC_Celerra_SIDS3/prodata/2/pubpna/US11D_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	397.6	99.6	504	US-09-925-065A-153932	Sequence 153932,
2	397.6	99.6	504	US-09-925-065A-153933	Sequence 153933,
3	397.6	99.6	504	US-09-925-065A-153932	Sequence 153932,
4	397.6	99.6	504	US-09-925-065A-153933	Sequence 153933,
5	397.6	99.6	509	US-10-301-480-247520	Sequence 247520,
6	397.6	99.6	509	US-10-301-480-247520	Sequence 247520,
7	397.6	99.6	509	US-10-301-480-247520	Sequence 247520,
8	397.6	99.6	509	US-10-301-480-247520	Sequence 247520,
9	397.6	99.6	509	US-10-301-480-247520	Sequence 247520,
10	397.6	99.6	509	US-10-301-480-247520	Sequence 247520,
11	397.6	99.6	509	US-10-301-480-247520	Sequence 247520,
12	397.6	99.6	509	US-10-301-480-247520	Sequence 247520,
13	397.6	99.6	509	US-10-301-480-247520	Sequence 247520,
14	397.6	99.6	509	US-10-301-480-247520	Sequence 247520,
15	397.6	99.6	509	US-10-301-480-247520	Sequence 247520,
16	397.6	99.6	509	US-10-301-480-247520	Sequence 247520,
17	397.6	99.6	509	US-10-301-480-247520	Sequence 247520,

18	54.2	13.6	599	US-09-925-065A-321660	Sequence 321660,
19	54.2	13.6	599	US-09-925-065A-321660	Sequence 321660,
20	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
21	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
22	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
23	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
24	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
25	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
26	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
27	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
28	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
29	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
30	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
31	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
32	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
33	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
34	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
35	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
36	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
37	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
38	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
39	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
40	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
41	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
42	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
43	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
44	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
45	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
46	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
47	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
48	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
49	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
50	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
51	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
52	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
53	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
54	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
55	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
56	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
57	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
58	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
59	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
60	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
61	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
62	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
63	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
64	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
65	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
66	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
67	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
68	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
69	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
70	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
71	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
72	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
73	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
74	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
75	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
76	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
77	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
78	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
79	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
80	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
81	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
82	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
83	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
84	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
85	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
86	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
87	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
88	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
89	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,
90	54.2	13.6	616	US-10-301-480-396282	Sequence 396282,

```

C 91 51 12.8 3464 8 US-10-763-210-26 Sequence 26, Appl
C 92 51 12.8 3467 8 US-10-763-210-25 Sequence 25, Appl
C 93 51 12.8 3467 8 US-10-763-210-29 Sequence 22, Appl
C 94 51 12.8 3469 8 US-10-763-210-27 Sequence 27, Appl
C 95 51 12.8 3470 8 US-10-763-210-28 Sequence 28, Appl
C 96 51 12.8 3473 8 US-10-763-210-24 Sequence 24, Appl
C 97 51 12.8 28854 3 US-09-741-151-3 Sequence 3, Appl1
C 98 50.8 12.7 591 4 US-09-925-065A-644552 Sequence 642552,
C 99 50.8 12.7 591 5 US-09-925-065A-642552 Sequence 642552,
C 100 50.8 12.7 644 12 US-10-301-480-391371 Sequence 391371,
C 101 50.8 12.7 644 12 US-10-301-480-1004780 Sequence 1004780,
C 102 50.8 12.7 962 12 US-10-301-480-544114 Sequence 544114,
C 103 50.8 12.7 962 12 US-10-301-480-1157523 Sequence 1157523,
C 104 50.8 12.7 92726 3 US-09-997-722-193 Sequence 193, App
C 105 50.8 12.7 194945 7 US-10-085-117-155 Sequence 355, App
C 106 50.6 12.7 531 4 US-09-925-065A-486389 Sequence 486389,
C 107 50.6 12.7 531 5 US-09-925-065A-486389 Sequence 486389,
C 108 50.6 12.7 562 12 US-10-301-480-294578 Sequence 294578,
C 109 50.6 12.7 562 12 US-10-301-480-907987 Sequence 907987,
C 110 50.6 12.7 593 4 US-09-925-065A-206820 Sequence 206820,
C 111 50.6 12.7 593 5 US-09-925-065A-206820 Sequence 206820,
C 112 50.6 12.7 633 4 US-09-925-065A-99858 Sequence 99858, A
C 113 50.6 12.7 633 5 US-09-925-065A-99858 Sequence 99858, A
C 114 50.6 12.7 640 12 US-10-301-480-200523 Sequence 200523,
C 115 50.6 12.7 640 12 US-10-301-480-813932 Sequence 813932,
C 116 50.6 12.7 694 6 US-10-027-632-111223 Sequence 111223,
C 117 50.6 12.7 694 7 US-10-027-632-111223 Sequence 111223,
C 118 50.4 12.6 612 4 US-09-925-065A-742315 Sequence 742315,
C 119 50.4 12.6 612 5 US-09-925-065A-742315 Sequence 742315,
C 120 50.4 12.6 640 12 US-10-301-480-71593 Sequence 71593, A
C 121 50.4 12.6 640 12 US-10-301-480-685002 Sequence 685002,
C 122 50.4 12.6 657 12 US-10-301-480-605379 Sequence 605379,
C 123 50.4 12.6 657 12 US-10-301-480-1218788 Sequence 1218788,
C 124 50.4 12.6 827 6 US-10-027-632-165881 Sequence 165881,
C 125 50.4 12.6 827 6 US-10-027-632-165882 Sequence 165882,
C 126 50.4 12.6 827 7 US-10-027-632-165881 Sequence 165881,
C 127 50.4 12.6 827 7 US-10-027-632-165882 Sequence 165882,
C 128 50.4 12.6 967 12 US-10-301-480-560457 Sequence 560457,
C 129 50.4 12.6 967 12 US-10-301-480-1173866 Sequence 1173866,
C 130 50.4 12.6 1000 12 US-10-301-480-557803 Sequence 557803,
C 131 50.4 12.6 1000 12 US-10-301-480-1171212 Sequence 1171212,
C 132 50.4 12.6 1994 4 US-09-925-065A-683118 Sequence 683118,
C 133 50.4 12.6 1994 5 US-09-925-065A-683118 Sequence 683118,
C 134 50.2 12.6 106 9 US-10-674-124A-25224 Sequence 25224, A
C 135 50.2 12.6 186 9 US-10-674-124A-24700 Sequence 24700, A
C 136 50.2 12.6 341 9 US-10-674-124A-11186 Sequence 11186, A
C 137 50.2 12.6 341 9 US-09-925-065A-156885 Sequence 156885,
C 138 50.2 12.6 379 5 US-09-925-065A-156885 Sequence 156885,
C 139 50.2 12.6 391 12 US-10-301-480-250077 Sequence 250077,
C 140 50.2 12.6 391 12 US-10-301-480-863486 Sequence 863486,
C 141 50.2 12.6 401 12 US-10-301-480-250076 Sequence 250076,
C 142 50.2 12.6 401 12 US-10-301-480-863485 Sequence 863485,
C 143 50.2 12.6 413 4 US-09-925-065A-156884 Sequence 156884,
C 144 50.2 12.6 413 4 US-09-925-065A-156884 Sequence 156884,
C 145 50.2 12.6 622 6 US-10-027-632-115412 Sequence 115412,
C 146 50.2 12.6 622 6 US-10-027-632-115412 Sequence 115412,
C 147 50.2 12.6 668 6 US-10-027-632-228805 Sequence 228805,
C 148 50.2 12.6 668 6 US-10-027-632-228805 Sequence 228805,
C 149 50.2 12.6 668 7 US-10-027-632-228805 Sequence 228805,
C 150 50.2 12.6 668 7 US-10-027-632-228806 Sequence 228806,

```

ALIGNMENTS

```

RESULT 1
US-09-925-065A-153932/c
; Sequence 153932, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome

```

```

; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243, 096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252, 147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250, 092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261, 766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289, 846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 153932
; LENGTH: 504
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-153932

Query Match 99.6%; Score 397.6; DB 4; Length 504;
Best Local Similarity 99.5%; Pred. No. 2.8e-102;
Matches 398; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAGATGCTCTTCTCTATATATTTAGAGTCTTGATAGCTTAATAATCTGTGGCG 60
DB 420 AAGATGCTCTTCTCTATATATTTAGAGTCTTGATAGCTTAATAATCTGTGGCG 361
QY 61 TACACATGACCTGCTGATCTGTTACTCTGTTGAAAGTGTGACGAGGAGTAAAG 120
DB 360 TACACATGAACTGCTGATCTGTTACTCTGTTGAAAGTGTGACGAGGAGTAAAG 301
QY 121 CTTACCTTTTGTACTGTTGATATATATATATATATATATATATATATATATATAT 180
DB 300 CTTACCTTTTGTACTGTTGATATATATATATATATATATATATATATATATATAT 241
QY 181 CAAATATTTTAACTCGGACATATATGATACACCACTTCAAGTAAAGTAACTACTATCT 240
DB 240 CAAATATTTTAACTCGGACATATATGATACACCACTTCAAGTAAAGTAAAGTAACT 181
QY 241 CAGAAATGAAACATATATCTGCTATTTGCTATTAATGACAGTAAAGTAAAGTAAAGT 300
DB 180 CAGAAATGAAACATATATCTGCTATTTGCTATTAATGACAGTAAAGTAAAGTAAAGT 121
QY 301 ATGAGTTCAGTTCCTTACTATGAAACCTGAAAGGTGAATCTGGTGTAACTTAACT 360
DB 120 ATGAGTTCAGTTCCTTACTATGAAACCTGAAAGGTGAATCTGGTGTAACTTAACT 61
QY 361 CTTTGAAGAACTGCAATACCTTAATCTCAAGATCAGACGAG 400
DB 60 CTTTGAAGAACTGCAATACCTTAATCTCAAGATCAGACGAG 21

```

```

RESULT 2
US-09-925-065A-153932/c
; Sequence 153932, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243, 096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252, 147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250, 092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261, 766
; PRIOR FILING DATE: 2001-01-16

```

Query Match	99.6%	Score 397.6	DB 5	Length 504
-------------	-------	-------------	------	------------

Oy	1	AAAAAGCCTTCTCTATAATTTACAGATCTGGAATAGCTTAAATACCTGTGGCG	60
Db	420	AAGAAAGCTCTTCTCTATAATTTACAGATCTTGAATAGCTTAAATACCTGTGGCG	361
Oy	61	TACACATGAGACTGTGATACCTGTACCTGGAAGTAGATGCAAGGCAAGTAGG	120
Db	360	TACACATGGAACGTGGATACCTGTACCTGTGGAAGTAGATGCAAGGCAAGTAGG	301

Oy		121	CCTACCTTTTGTAGCGTTGATATTATTAATAAACAACAACAACAACAACAACA	180
Db		300	CCTACTTTTTTGTCAGTGGTATATATATATATATATATATATATATATATATAT	241
Oy		181	CAVATATTTAACTCGGAACATMAATGTACACCACCTATTCAAAGAATAAAACTAATCT	240
Db		240	CACATATTTTAACTCGGAACATMAATGTACACCACCTATTCAAAGAATAAAACTAATCT	181
Oy		241	CAGAAGTAAAGCATATATCCTGTATATGCATTAACTAGACAGTAAAGCTGCCAACCTT	300
Db		180	CAGAAGTAAAGCATATATCCTGTATATGCATTAACTAGACAGTAAAGCTGCCAACCTT	121
Oy		301	ATGCAGTTCAGTTCCCTACTATGAAAACCTGAAAGGTGAAAATCTGGNGSTTTACCTAA	366
Db		120	ATGCAGTTCAGTTCCCTACTATGAAAACCTGAAAGGTGAAAATCTGGNGSTTTACCTAA	61
Oy		361	CCTTGAAAACGTCATTAACCTAATCTAGATCAAGACGGAG +40	
Db		60	CCTTGAAAACGTCATTAACCTAATCTAGATCAAGACGGAG 21	

RESULT 5
US-10-301-480-247519/c
; Sequence 247519, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:

```

1  APPLICANT: Wang, David G.
2  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
3  TITLE OF INVENTION: in the Human Genome
4  FILE REFERENCE: 108827.137
5  CURRENT APPLICATION NUMBER: US/10/301,480
6  CURRENT FILING DATE: 2002-11-21
7  PRIOR APPLICATION NUMBER: US 10/215,598
8  PRIOR FILING DATE: 2002-08-09
9  PRIOR APPLICATION NUMBER: US 60/311,635
10 PRIOR FILING DATE: 2001-08-10
11 NUMBER OF SEQ ID NOS: 1226818
12 SOFTWARE: FastSeq for Windows Version 4.0
13 SEQ ID NO 247519
14 LENGTH: 509
15 TYPE: DNA
16 ORGANISM: Homo sapien
17 US-10-301-480-247519

```

Query Match	99.6%	Score 397.6	DB 12	Length 509
Best Local Similarity	99.5%	Pred. No. 2.8e-102		
Matches 398	Conservative	1	Mismatches	0
			Indels	0
			Gaps	0

QY	1	AAGAATGCTCTTCTCTATATATTACGAATCTTGAATAGCTTAATAATCTGTGGC	60
Db	425	AAGAATGCTCTTCTCTATATATTACGAATCTTGAATAGCTTAATAATCTGTGGC	366
QY	61	TACACATGAGACTGSGTATCTGGTTACCTCTGGAAGTGTGCAAGGCAAGTGAAG	120
Db	365	TACACATGGAACGTGTACTGTGTTACCTGTGAAGGTGTATGCAAGGCAAGTGAAG	306
QY	121	CCTACCTTTTGGTACTGTTGATATATATATATACATACACACACACACACACA	180
Db	305	CCTACCTTTTGGTACTGTTGATATATATATATATACATACACACACACACACACA	246
QY	181	CAVATATTTTAACTGGGACATTAATGTACACACCTTATTCAAAGTAAAACTACTT	240
Db	245	CACATATTTTAACTGGGACATTAATGTACACACCTTATTCAAAGTAAAACTACTT	186
QY	241	CAGAAATTAAGAATATAATCTCTGCTATTGTCAATTACTAGAGATTAACCTGCTAACTTT	300
Db	185	CAGAAATTAAGAATATAATCTCTGCTATTGTCAATTACTAGAGATTAACCTGCTAACTTT	120
QY	301	ATGCAGTTCAAGTTCCCTCACTATGAAAACTGAAAGGTTAAATCTGTGTTTAACTTAA	366
Db	125	ATGCAGTTCAAGTTCCCTCACTATGAAAACTGAAAGGTTAAATCTGTGTTTAACTTAA	66
QY	361	CCTTTGAAAAAGTCAATATCTTAACTCAGATCAGAACGAG	400

Db 65 CCTTGAACGTCATACCTACTCAGATCAGAAGGAG 26

```

RESULT 6
US-10-301-480-247520/c
; Sequence 247520, Application US/10301480
; Publication NO. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 122618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 247520
; LENGTH: 509
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-247520

```

Query Match	. 99.6%	Score 397.6;	DB 12;	Length 509;
Best Local Similarity	99.2%;	Pred. No. 2.8e-102;		
Matches 397;	Conservative 2;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	AAGAAAGCCTTCTCTCTATATATTTAGAGGTTCTTGAATAGAGTTAAATATCCGTGGCG	60
Db	425	AAGAAAGCCTTCTCTATATATTTAGAGGTTCTTGAATAGAGTTAAATATCCGTGGCG	366
QY	61	TACACATGAGACTGATGATCTGTTACCTCTGGAAAGTAGATGCAAGGCAAGTGAGG	120
Db	365	TACACATGAACTGGTGATCTAGGTTACCTCTGGAAAGTAGATGTCAGGGCAAGTGAGG	306
QY	121	CCTACCTTTTGTACTGTTGATATATATATATACATACACACACACACACACACA	180
Db	305	CCTACCTTTTGTACTGTTGATATATATATATATACATACACACACACACACACACA	246
QY	161	CATATATTTTAACTCGGGACATTAAGTACACACACCATTCAAAGTAAAACTACTACT	240
Db	245	CACATATTTTAACTCGGGACATTAAGTACACACACCATTCCTTCAAGTAAAACTACTACT	186
QY	241	CAGAAGTAAAGACATATATCTGCTATATGTCATTAACTAGCAGATTAAAGCTGCTAACCTTT	300
Db	185	CAGAAGTAAAGACATATCTGCTATATGTCATTAACTAGCAGATTAAAGCTGCTAACCTTT	126
QY	301	ATGCAGTTCAGTTTCCCTACTATATGAAAACTGAAAAGTTGAAATCTGTGCTTTAACTTAA	360
Db	125	ATGCAGTTCAGTTTCCCTACTATATGAAAACTGAAAAGTTGAAATCTGTGCTTTAACTTAA	66
QY	361	CCTTTGAAAACGTCATACCTTAACTAGATCAGAACGAGG	400
Db	65	CCTTTGAAAACGTCATACCTTAACTAGATCAGAACGAGG	26

```

RESULT 7
US-10-301-480-860928/C
; Sequence 860928, Application US/10301480
; Publication No. US2006005756A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: In the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09

```



```

? PRIOR APPLICATION NUMBER: US 60/311,695
? PRIOR FILING DATE: 2001-08-10
? NUMBER OF SEQ ID NOS: 1226818
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 860928
?
? LENGTH: 509
? TYPE: DNA
? ORGANISM: Homo sapien
? US-10-301-480-860928

```

Query Match	99.6%;	Score 397.6;	DB 12;	Length 509;
Best Local Similarity	99.5%;	Pred. No. 2.8e-102;		
Matches 398; Conservative	1;	Mismatches 1;	Indels 0;	Gaps 0;

[illegible]

```

RESULT 8
US-10-301-480-860929/C
; Sequence 860929, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome

```

```

? CURRENT FILING DATE: 2002-11-21
? CURRENT APPLICATION NUMBER: US/10/301,480
? PRIOR FILING DATE: 2002-08-09
? PRIOR APPLICATION NUMBER: US 60/311,695
? PRIOR FILING DATE: 2001-08-10
? NUMBER OF SEQ ID NOS: 1226816
? SOFTWARE: FASTSEQ for Windows Version 4.0.
? SEQ ID NO 860929
? LENGTH: 509
? TYPE: DNA
? ORGANISM: Homo sapien
US-10-301-480-860929

```

Query Match	99.6%	Score 397.6;	DB 12;	Length 509;
Best Local Similarity	99.2%	Pred. No. 2.8e-102;		
Matches 397; Conservative	2;	Mismatches 1;	Indels 0;	Gaps 0;

1 AAGAAGCTCTTCTCTATATATTACGAGTCTTGATAAGCTTAATAATACCTGTGCG 60
425 AAGAAGCTCTTCTCTATATATTACGAGTCTTGATAAGCTTAATAATACCTGTGCG 366

Qy	61	TACACATGAGCTGTGATCTGCTTACTCGTAAAGTAGTAATGCAGGGCAAGGAG	120
Dp	365	TACACATGAACCTGGTGTACTGGTACTCCTCGAAAGTAGTAATGCAGGGCAAGGAG	306
Qy	121	CCTACCTTTTGTACTGTTGTATATATATATATATACATACACACACACACACACAC	180
Dp	305	CCTACCTTTTGTACTGTTGTATATATATATATATATACATACACACACACACACAC	246
Qy	181	CATATATTTTAACCTGGGACATAAATGTACACACACCTATTCAAAGTAAAACTACTATCT	240
Dp	245	CACATATTTTAACTCGGGACATAAATGTACACACCTATTCAAAGTAAAACTACTATCT	186
Qy	241	CAGAAGTAAAGCATATATCTCGTATTTGTCAATTACTGACAGTTAAGCTGCTTAACCTT	300
Dp	185	CAGAAGTAAAGCATATATCTCGTATTTGTCAATTACTGACAGTTAAGCTGCTTAACCTT	126
Qy	301	ATGAGATTCAATTTCCCTACTATATGAAACCTGAAGGTTGAATCTGAGTGTTAACCTAA	360
Dp	125	ATGAGATTCAATTTCCCTACTATATGAAACCTGAAGGTTGAATCTGAGTGTTAACCTAA	66
Qy	361	CCTTGAAAAAGTCAATACCTTAACCTCAGATCAGAAAGGAG	400
Dp	65	CCTTGAAAAAGTCAATACCTTAACCTCAGATCAGAAAGGAG	26

RESULT 9
US-10-723-518-1
; Sequence 1, Application US/10723518
; Publication No. US20050064442A1

```

1  APPLICANT: ROTH, RICHARD B.
2  APPLICANT: NELSON, MATTHEW ROBERTS
3  APPLICANT: BRAUN, ANDREAS
4  APPLICANT: KAMERER, STEFAN M.
5  APPLICANT: DENISENKO, MIKHAIL F.
6  APPLICANT: RENELAND, RIKARD
7  APPLICANT: ATIENZA, JOSEPHINE M.
8  TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF BREAST CANCER AND TREATMENTS
9  TITLE OF INVENTION: THEREOF
10 FILE REFERENCE: SEQ-4068-UT
11 CURRENT APPLICATION NUMBER: US/10/723,518
12 CURRENT FILING DATE: 2003-11-25
13 PRIOR APPLICATION NUMBER: US 60/429,136
14 PRIOR FILING DATE: 2002-11-25
15 PRIOR APPLICATION NUMBER: US 60/490,234
16 PRIOR FILING DATE: 2003-07-24
17 PRIOR APPLICATION NUMBER: US 60/504,258
18 PRIOR FILING DATE: 2003-09-18
19 NUMBER OF SEQ ID NOS: 235
20 SOFTWARE: PatentIn version 3.2
21 SEQ ID NO 1
22 LENGTH: 83405
23 TYPE: DNA
24 ORGANISM: Homo sapiens
25 FEATURE:
26 NAME/KEY: misc feature
27 LOCATION: (18828)..(18833)
28 OTHER INFORMATION: this region may or may not be present
29 US-10-723-518-1

```

Query Match	99.6%	Score 397.6;	DB 10;	Length 83405;
Best Local Similarity	99.8%	Pred. No. 2.6e-101;		
Matches 399; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	AAAGAAAGCCTTTCCTATATATATATTAACAGGCTCTTGAATTAACCTTAAATACCTGTGGG	60
Db	33601	AAAGAAAGCCTTTCCTATATATATATTAACAGGCTCTTGAATTAACCTTAAATACCTGTGGG	33660
QY	61	TACACATGAGACTGTGTACTGCTTACTCTGGAAGTAGGTATGACGGCAAGTAGG	120
Db	33661	TACACATGGAACGTGTGTACTGCTTACTCTGGAAGTAGGTATGACGGCAAGTAGG	33720


```
Publication No. US20060057564A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1151147
LENGTH: 707
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-1151147

Query Match      14.2%; Score 56.6; DB 12; Length 707;
Best Local Similarity 64.3%; Pred. No. 2.5e-05;
Matches 83; Conservative 1; Mismatches 45; Indels 0; Gaps 0;

QY 122 CTACCTTTTGTACTGTTGATATATATATATATACATACACACACACACACACAC 181
DB 146 CAACATATATGCAATTTCTCATATATATACACACACACACACACACACACACAC 205
QY 182 AATATTTTAACTCGGACATATAATGTACACACCTTATTCAAATGTAATACTATCTC 241
DB 206 AATATTTTCAATCATATATATATTTGGCCACACCTGAAAGGAATAGAAATGAAGA 265
QY 242 AGAAGTAAA 250
DB 266 AGAAGATA 274

RESULT 17
US-09-918-995-11490/C
Sequence 11490, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11490
LENGTH: 463
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)...(463)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-11490

Query Match      14.0%; Score 56; DB 3; Length 463;
Best Local Similarity 62.3%; Pred. No. 3e-05;
Matches 86; Conservative 1; Mismatches 51; Indels 0; Gaps 0;
```

```
QY 226 TAAAACTACTATCTCAG 243
DB 223 AATCAAAAAATGTGAAG 206

RESULT 18
US-09-925-065A-321660
Sequence 321660, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 321660
LENGTH: 599
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-321660

Query Match      13.6%; Score 54.2; DB 4; Length 599;
Best Local Similarity 66.7%; Pred. No. 0.00011;
Matches 78; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 76 TGATACGTGTTACCTCTGGAAGTAGTACGAGGGGCAAGTGGCTTACCTTTTGTAC 135
DB 142 TGAATAATCTTATTTTCAAAAAATGTCATGGAAGTAGCAAGATTCATTAATAAA 201
QY 136 TGTGTATATATATATATATACATACACACACACACACACACACACACATATTTAA 192
DB 202 TCAATATATATATATATATATATATACACACACACACACACACACACATATATTTA 258

RESULT 19
US-09-925-065A-321660
Sequence 321660, Application US/09925065A
Publication No. US20050228172A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 321660
LENGTH: 599
```

TYPE: DNA
ORGANISM: Homo sapiens
US-10-301-480-321660

Query Match 13.6%; Score 54.2; DB 5; Length 599;
Best Local Similarity 66.7%; Pred. No. 0.00011;
Matches 78; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 76 TGAATCTGTTACCTCTGGAAGTGTATGACGGGCAAGTGAAGCTTCTTTTGTAC 135
DB 142 TGAATAATCTTATTTTCAAAAATTTGTCATGAGTAGTGCAGATTCATTAAAAA 201
QY 136 TGTGTATATATATATATATATACACACACACACACACACACATATATTTAA 192
DB 202 TCAATATATATATATATATATATATATACACACACACACACACACATATATTTA 258

RESULT 20

US-10-301-480-396282
Sequence 396282, Application US/10301480
Publication No. US20060057564A1

GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 396282
LENGTH: 616
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-396282

Query Match 13.6%; Score 54.2; DB 12; Length 616;
Best Local Similarity 66.7%; Pred. No. 0.00011;
Matches 78; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 76 TGAATCTGTTACCTCTGGAAGTGTATGACGGGCAAGTGAAGCTTCTTTTGTAC 135
DB 159 TGAATAATCTTATTTTCAAAAATTTGTCATGAGTAGTGCAGATTCATTAAAAA 218
QY 136 TGTGTATATATATATATATATACACACACACACACACACACATATATTTAA 192
DB 219 TCAATATATATATATATATATATATATACACACACACACACACATATATTTA 275

RESULT 21

US-10-301-480-1009691
Sequence 1009691, Application US/10301480
Publication No. US20060057564A1

GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1009691
LENGTH: 616
TYPE: DNA

ORGANISM: Homo sapien
US-10-301-480-1009691

Query Match 13.6%; Score 54.2; DB 12; Length 616;
Best Local Similarity 66.7%; Pred. No. 0.00011;
Matches 78; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 76 TGAATCTGTTACCTCTGGAAGTGTATGACGGGCAAGTGAAGCTTCTTTTGTAC 135
DB 159 TGAATAATCTTATTTTCAAAAATTTGTCATGAGTAGTGCAGATTCATTAAAAA 218
QY 136 TGTGTATATATATATATATATACACACACACACACACACACATATATTTAA 192
DB 219 TCAATATATATATATATATATATATATACACACACACACACACATATATTTA 275

RESULT 22

US-10-301-480-566528/c
Sequence 566528, Application US/10301480
Publication No. US20060057564A1

GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 566528
LENGTH: 682
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-566528

Query Match 13.5%; Score 54; DB 12; Length 682;
Best Local Similarity 59.2%; Pred. No. 0.00013;
Matches 90; Conservative 1; Mismatches 61; Indels 0; Gaps 0;

QY 113 AAGTAGGCTTACCTTTTGTACTGTTGATATATATATATATATACACACACACAC 172
DB 525 ACGGTAGGCAATGATGTGTGTCAGTTAAAAAGTTTCTACACACACACACACAC 466
QY 173 CACACACACATATTTTAACTCGGACATTAATGTACACACACCTATTCAAGTAAAC 232
DB 465 CACACACACACACACACACACATTTCTTAAGATTTACACACATATCATGAAATCC 406
QY 233 TACTATCTCAGAGTAAGATTAAGCATATATCTGCT 264
DB 405 TTCTACTGAGTAGCAGATATATAGCAAACT 374

RESULT 23

US-10-301-480-1179937/c
Sequence 1179937, Application US/10301480
Publication No. US20060057564A1

GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 1179937
; LENGTH: 682
; TYPE: DNA
; ORGANISM: Homo sapien
; US-10-301-480-1179937

```

Query Match	13.5%	Score 54;	DB 12;	Length 682;
Best Local Similarity	59.2%	Pred. No. 0.00013;		
Matches	90;	Conservative	1;	Mismatches 61;
			Indels	0;
			Gaps	0;

QY	DB	QY	DB
113	525	173	465
AAATGAGGCGTACCTTTTGTACTGTTTGATTTATATATATATACACACACACA	ACCGTGGGATAGTATGTGTCCAGTTAAAAAAGTTTCAACACACACACACACA	CACACACACATATTTTAACTCGGGACATAATATACACCACCTATTCAAAGTAAAC	CACACACACACACACACACACACATTCCTAAAGATTACACAAATCATCATGATCC
264	374	233	405
TACTATCTCAAGATTAAGACATTAATCCGCT	TTCTAACTAGTAGCAAGATTATGACCAACT		

RESULT 24
US-10-674-124A-20148
; Sequence 20148, Application US/10674124A
; Publication No. US20040197797A1

```

1  APPLICANT: INOKO, Hideo
2  APPLICANT: TAMUYA, Gen
3  TITLE OF INVENTION: GENE MAPPING METHOD USING MICROATELLE
4  TITLE OF INVENTION: GENETIC POLYMORPHISM MARKERS
5  FILE REFERENCE: ORIN-003CIP
6  CURRENT APPLICATION NUMBER: US/10/674,124A
7  CURRENT FILING DATE: 2003-09-26
8  PRIOR APPLICATION NUMBER: 10/257,511
9  PRIOR FILING DATE: 2003-03-07
10 PRIOR APPLICATION NUMBER: PCT/JP00/07621
11 PRIOR FILING DATE: 2000-10-30
12 PRIOR APPLICATION NUMBER: JP2000-112659
13 PRIOR FILING DATE: 2000-04-13
14 PRIOR APPLICATION NUMBER: JP2002-327516
15 PRIOR FILING DATE: 2002-09-28
16 PRIOR APPLICATION NUMBER: JP2002-383869
17 PRIOR FILING DATE: 2002-12-09
18 NUMBER OF SEQ ID NOS: 27110
19 SEQ ID NO 20148
20
21 LENGTH: 372
22 TYPE: DNA

```

OTHER INFORMATION: chr13.fa.07fzr.84832263
FEATURE:
OTHER INFORMATION: Located on chromosome 13
FEATURE:
OTHER INFORMATION: Distance between a terminus base of telomere on
OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
OTHER INFORMATION: sequence : 82516891
FEATURE:
OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
OTHER INFORMATION: 5'-terminus of this base sequence : 172757
DS-10-674-124A-20148

Query Match	13.5%	Score 53.8;	DB 9;	Length 372;
Best Local Similarity	64.2%;	Pred. No. 0.00012;		
Matches	79;	Conservative	1;	Mismatches 43;
			Indels	0;
			Gaps	0

OY 134 ACCTTTGAAATATATATATACAGTACACACACACACAATATTTTAAC 19

Dd 120 AATGTGTTTGATATATATAATATACACACACACACAATATATATGTTT 175

OY 194 TCGGGACATATAATGTCACCACCTATTCCAAAGTAAAACTACTATTCTGAAGATTTAAGC 255

Db 180 CCATTGTGATATATATAGTTAGCCTTAAAAAGTAGAAATTCGTGATTTGAGTAAAC 239
 QY 254 ATA 256
 Db 240 AGA 242

RESULT 25
US-09-925-065A-646805

```

1  APPLICANT: Wang, David G.
2  TITLE OF INVENTION: Identification and Mapping of Single
3  TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
4  FILE REFERENCE: 108827.135
5  CURRENT APPLICATION NUMBER: US/09/925,065A
6  CURRENT FILING DATE: 2001-08-08
7  PRIOR APPLICATION NUMBER: US 60/243,096
8  PRIOR FILING DATE: 2000-10-24
9  PRIOR APPLICATION NUMBER: US 60/252,147
10 PRIOR FILING DATE: 2000-11-20
11 PRIOR APPLICATION NUMBER: US 60/250,092
12 PRIOR FILING DATE: 2000-11-30
13 PRIOR APPLICATION NUMBER: US 60/261,766
14 PRIOR FILING DATE: 2001-01-16
15 PRIOR APPLICATION NUMBER: US 60/289,846
16 PRIOR FILING DATE: 2001-05-09
17 NUMBER OF SEQ ID NOS: 957086
18 SOFTWARE: FastSeq for Windows Version 4.0
19 SEQ ID NO 646805
20 LENGTH: 588
21 TYPE: DNA
22 ORGANISM: Homo sapiens
23 US-09-925-065A-646805

```

Query Match	13.5%	Score 53.8;	DB 4;	Length 588;
Best Local Similarity	67.0%	Pred. No. 0.00014;		
Matches 77;	Conservative 0;	Mismatches 38;	Indels 0;	Gaps 0

[illegible]

RESULT 26
US-09-925-065A-646805
; Sequence 646805, Application US/09925065A
; Publication No. US20050228172A9

```

1  APPLICANT: Wang, David G.
2  TITLE OF INVENTION: Identification and Mapping of Single
3  TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
4  FILE REFERENCE: 106827.135
5  CURRENT APPLICATION NUMBER: US/09/925.065A
6  CURRENT FILING DATE: 2001-08-08
7  PRIOR APPLICATION NUMBER: US 60/243,096
8  PRIOR FILING DATE: 2000-10-24
9  PRIOR APPLICATION NUMBER: US 60/252,147
10 PRIOR FILING DATE: 2000-11-20
11 PRIOR APPLICATION NUMBER: US 60/250,092
12 PRIOR FILING DATE: 2000-11-30
13 PRIOR APPLICATION NUMBER: US 60/261,766
14 PRIOR FILING DATE: 2001-01-16
15 PRIOR APPLICATION NUMBER: US 60/289,846
16 PRIOR FILING DATE: 2001-05-09
17 NUMBER OF SEQ ID NOS: 957086
18 SOFTWARE: FastSeq for Windows Version 4.0

```


GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 132528
LENGTH: 578
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-132528

Query Match 13.3%; Score 53.2; DB 12; Length 578;
Best Local Similarity 58.7%; Pred. No. 0.00021;
Matches 88; Conservative 2; Mismatches 60; Indels 0; Gaps 0;

95 AAAGTAGGTATGCGAGGCGAAGTGGCTTACTCTTTTGACTGTTGATATATATATATA 154
150 ATATTATATATACACAGATTTTATATATATATATATATATATATATATATATA 91
155 CATACCA 214
90 TAT 31
215 CCTATTCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 244
30 CAAAACAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1

RESULT 34

US-10-301-480-745936/c
Sequence 745936, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 745936
LENGTH: 578
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-745936

Query Match 13.3%; Score 53.2; DB 12; Length 578;
Best Local Similarity 58.7%; Pred. No. 0.00021;
Matches 88; Conservative 2; Mismatches 60; Indels 0; Gaps 0;

95 AAAGTAGGTATGCGAGGCGAAGTGGCTTACTCTTTTGACTGTTGATATATATATATA 154
150 ATATTATATATACACAGATTTTATATATATATATATATATATATATATATATA 91
155 CATACCA 214
90 TAT 31
215 CCTATTCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 244

Db 30 CAAAACAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1

RESULT 35
US-10-301-480-745937/c
Sequence 745937, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 745937
LENGTH: 578
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-745937

Query Match 13.3%; Score 53.2; DB 12; Length 578;
Best Local Similarity 58.7%; Pred. No. 0.00021;
Matches 88; Conservative 2; Mismatches 60; Indels 0; Gaps 0;

95 AAAGTAGGTATGCGAGGCGAAGTGGCTTACTCTTTTGACTGTTGATATATATATATA 154
150 ATATTATATATACACAGATTTTATATATATATATATATATATATATATATATA 91
155 CATACCA 214
90 TAT 31
215 CCTATTCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 244
30 CAAAACAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1

RESULT 36

US-10-301-480-411688/c
Sequence 411688, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 411688
LENGTH: 535
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-411688

Query Match 13.3%; Score 53; DB 12; Length 535;
Best Local Similarity 61.5%; Pred. No. 0.00023;
Matches 83; Conservative 1; Mismatches 51; Indels 0; Gaps 0;

111 GCAAGTAGGCTTACTCTTTTGACTGTTGATATATATATATATATATATATATATAT 170
236 GAAAGAGGCCCATCTCTGCTGCTGATATATATATATATATATATATATATATAT 177

QY	Db	QY	Db
171	CAACACACACAAATATTTTAACTGGGACATPAATTTAGACCAACCTTTCAAAAGTAAA	231	ACTKCTATCTCGAAA 245
176	CACACACACACACATATGTATGTGTGTATATGTATATATACACCAACAACTATATAG	116	GCCTGCCCTTTTAAAA 102

RESULT 37
US-10-301

US-10-301-480-411689/c
; Sequence 411689, Application US/10301480
; Publication No. US20060057564A1
GENERAL INFORMATION:

```

1  APPLICANT: Wang, David G.
2  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
3  TITLE OF INVENTION: in the Human Genome
4  FILE REFERENCE: 108827.137
5  CURRENT APPLICATION NUMBER: US/10/301,480
6  CURRENT FILING DATE: 2002-11-21
7  PRIOR APPLICATION NUMBER: US 10/215,598
8  PRIOR FILING DATE: 2002-08-09
9  PRIOR APPLICATION NUMBER: US 60/311,695
10 PRIOR FILING DATE: 2001-08-10
11 NUMBER OF SEQ ID NOS: 1226818
12 SOFTWARE: FastSeq for Windows Version 4.0
13 SEQ ID NO 411689
14 LENGTH: 535
15 TYPE: DNA
16 ORGANISM: Homo sapien
17 US-10-301-480-411689

```

```
Query Match      13.3%;   Score 53;   DB 12;   Length 535;
Best Local Similarity 61.5%;   Pred No. 0.00023;
Matches      83;   Conservative      1;   Mismatches      51;   Indels      0;   Gaps      0;
```

Oy	111	GCAGATGAGGCCGCTACCTTTTGGACGTGTGATATATATATATACATACACACACACA	170
Db	236	GGAAGAGGCCCATCTCTTGCTTGCTGATATATATATATATATACACACACACACACA	177
Oy	171	CACACACACACATATTTTAACTGGGGAATATATGTACACCACTATTCAAAGTAAAA	230
Db	176	CACACACACACACATATATATATATGTGTTGTGTATATATATATACACAAACAACTATAG	117
Oy	231	ACTACTATCTCAGAA	245
Db	116	GCTTGCTTTTAAAA	102

RESULT 38
US-10-301

US-10-301-480-1025097/c
; Sequence 1025097, Application US/10301480
; Publication No. US20060057564A1
; ~~US-10-301-480-1025097/c~~

```

1  APPLICANT: Wang, David G.
2  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
3  TITLE OF INVENTION: in the Human Genome
4  FILE REFERENCE: 108827.137
5  CURRENT APPLICATION NUMBER: US/10/301,480
6  CURRENT FILING DATE: 2002-11-21
7  PRIOR APPLICATION NUMBER: US 10/215,558
8  PRIOR FILING DATE: 2002-08-09
9  PRIOR APPLICATION NUMBER: US 60/311,695
10 PRIOR FILING DATE: 2001-08-10
11 NUMBER OF SEQ ID NOS: 1226818
12 SOFTWARE: FastSeq for windows Version 4.0
13 SEQ ID NO 1025097

```

Query Match	13.3%	Score 53	DB 12	Length 535
Best Local Similarity	61.5%	Pred. No. 0.00023		
Matches	83	Conservative	1	Mismatches 51
				Indels 0
				Gaps 0

QY	111	GCAAGGAGAGCCCTACCTTTTGTACTGTTGATATATATATATATACATACACACACACA	170
Db	236	GGAGAGGGCCCATCTTCTTGCTGTGTATATATATATATATATATACACACACACACA	177
QY	171	CACACACACACATATTTTAACTCGGACATTAATGTACACCACTTATCAAAAGTAAAA	230
Db	176	CACACACACACACATATGTATGTGTGTATATATATATATACACAAACATCTATAG	117
QY	231	ACTACTATCTCAGAA	245
Db	116	GCTTGCCTTTAAAA	102

RESULT 39

US-10-301-480-1025098/c
; Sequence 1025098, Application US/10301480
; Publication No. US20060057564A1
CENTRAL INTELLIGENCE AGENCY

```

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
TITLE OF INVENTION: In the Human Genome
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 122618
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1025098
LENGTH: 535
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-1025098

```

Query Match	13.3%	Score 53;	DB 12;	Length 535;
Best Local Similarity	61.5%;	Pred. No. 0.00023;		
Matches	83;	Conservative	1;	Mismatches 51; Indels 0; Gaps 0

Qy	111	GCAAGGAGGCCACCTTTTGTACTGTTTGATATATATATATACATACACACACACA	170
Db	236	GGAAGAGGCCCATCTCTTGCTGCTGATATATATATATATATATACACACACACACA	177
Qy	171	CACACACACACATATTTTAACTCGAGGACATAAATGTACACCACTTTCAAAGTAAA	230
Db	176	CACACACACACACATATATATATGTGTGTATATATATATATACACACACATCTATAG	117
Qy	231	ACTACTATCTCAGAA	245
Db	116	GCTTGCTTTTAAAA	102

RESULT 4C
US-10-995

US-10-995-561-26459/c
; Sequence 26459, Application US/10995561
; Publication No. US20050272054A1

```

1  APPLICANT: CARGILL, Michele et al.
2  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
3  TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
4  TITLE OF INVENTION: DETECTION AND USES THEREOF
5  FILE REFERENCE: CL001559
6  CURRENT APPLICATION NUMBER: US/10/995,561
7  CURRENT FILING DATE: 2004-11-24
8  NUMBER OF SEQ ID NOS: 85702
9  SOFTWARE: FASTSEQ for Windows Version 4.0
10 SEQ ID NO 25459
11 LENGTH: 201

```

[illegible]

```

OY      262  GCTATGTCATTAAGT 279
         ||| ||| ||| |||
Db      64. GCAGATTCTTAAAAAG 47

RESULT 42
US-09-925-065A-664473/C
/ Sequence 664473, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 664473
/ LENGTH: 504
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-664473

Query Match      13.2%; Score 52.6; DB 4; Length 504;
Beet Local Similarity 59.4%; Pred.No. 0.00029;
Matches 85; Conservative 2; Mismatches 56; Indels 0; Gaps 0

OY      53  CTGTGGCTCAACATGAGACTGSGTATCTGGTTACCTCTGGAAAGTAGTACGAGGC 112
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      439  CTGTCAATTCAGATGATGTCCTGATTTTCTGAAATTAACCTTCGTCACTGGTTTGTGGGC 380

OY      113  AAGTAGGCGCTACCTTTTGTACTGTTGATATATATATATATATACATACACACACACA 172
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      379  CAGTGCAGGCACTGTACTGTATTTGCAAAACACACACACACACACACACACACACA 320

OY      173  CACACACACAYATATTTTACTC 195
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      319  CACACACACACACWCCTCTCTC 297

RESULT 43
US-09-925-065A-664474/C
/ Sequence 664474, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09

```

NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 664474
LENGTH: 504
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-664474

Query Match 13.2%; Score 52.6; DB 5; Length 504;
Best Local Similarity 59.4%; Pred. No. 0.00029;
Matches 85; Conservative 2; Mismatches 56; Indels 0; Gaps 0;

QY 53 CTGTGGCGGTACATGAGTGTGATCTGTTACCTCTGGAAGTAGTATGACAGGC 112
DB 439 CTGTGCATCTCAGATGTCCTGATTTTCTGAATTACCTTGTCTCACTGTTGTGGCC 380
QY 113 AAGTAGGCGCTACCTTTTGTACTGTTGATATATATATATATACACACACACA 172
DB 379 CAGTGAAGCCACTGTAAGTATGCAAAACACACACACACACACACACACACA 320
QY 173 CACACACACAYATTTTAACTC 195
DB 319 CACACACACACACTCTCTCTC 297

RESULT 44

US-09-925-065A-664473/c
Sequence 664473, Application US/09925065A
Publication No. US20050228172A9
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925, 065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243, 096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252, 147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250, 092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261, 766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289, 846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 664473
LENGTH: 504
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-664473

Query Match 13.2%; Score 52.6; DB 5; Length 504;
Best Local Similarity 59.4%; Pred. No. 0.00029;
Matches 85; Conservative 2; Mismatches 56; Indels 0; Gaps 0;

QY 53 CTGTGGCGGTACATGAGTGTGATCTGTTACCTCTGGAAGTAGTATGACAGGC 112
DB 439 CTGTGCATCTCAGATGTCCTGATTTTCTGAATTACCTTGTCTCACTGTTGTGGCC 380
QY 113 AAGTAGGCGCTACCTTTTGTACTGTTGATATATATATATATACACACACACA 172
DB 379 CAGTGAAGCCACTGTAAGTATGCAAAACACACACACACACACACACACACA 320
QY 173 CACACACACAYATTTTAACTC 195
DB 319 CACACACACACACTCTCTCTC 297

RESULT 45
US-09-925-065A-664474/c

Sequence 664474, Application US/09925065A
Publication No. US20050228172A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925, 065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243, 096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252, 147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250, 092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261, 766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289, 846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 664474
LENGTH: 504
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-664474

Query Match 13.2%; Score 52.6; DB 5; Length 504;
Best Local Similarity 59.4%; Pred. No. 0.00029;
Matches 85; Conservative 2; Mismatches 56; Indels 0; Gaps 0;

QY 53 CTGTGGCGGTACATGAGTGTGATCTGTTACCTCTGGAAGTAGTATGACAGGC 112
DB 439 CTGTGCATCTCAGATGTCCTGATTTTCTGAATTACCTTGTCTCACTGTTGTGGCC 380
QY 113 AAGTAGGCGCTACCTTTTGTACTGTTGATATATATATATATACACACACACA 172
DB 379 CAGTGAAGCCACTGTAAGTATGCAAAACACACACACACACACACACACACA 320
QY 173 CACACACACAYATTTTAACTC 195
DB 319 CACACACACACACTCTCTCTC 297

RESULT 46

US-09-925-065A-646420/c
Sequence 646420, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925, 065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243, 096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252, 147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250, 092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261, 766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289, 846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 646420
LENGTH: 601
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-646420

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OW nucleic - nucleic search, using sw model

Run on: July 17, 2006, 16:33:23 ; Search time 259 Seconds
(without alignments)
2153.288 Million cell updates/sec

Title: SEQ1-33670G
Perfect score: 399.2
Sequence: 1.aagatgctctctctctccta.....taactcagatcagaagagag 400

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 886355 seqs, 697127050 residues

Total number of hits satisfying chosen parameters: 1772710

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 150 summaries

Database : Published Applications NA New:

1: /EMC_Celerra_SIDS3/ptocdata/1/pubpna/US09_NEW_PUB_seq.*
2: /EMC_Celerra_SIDS3/ptocdata/1/pubpna/US06_NEW_PUB_seq.*
3: /EMC_Celerra_SIDS3/ptocdata/1/pubpna/US07_NEW_PUB_seq.*
4: /EMC_Celerra_SIDS3/ptocdata/1/pubpna/US08_NEW_PUB_seq.*
5: /EMC_Celerra_SIDS3/ptocdata/1/pubpna/US09_NEW_PUB_seq.*
6: /EMC_Celerra_SIDS3/ptocdata/1/pubpna/US10_NEW_PUB_seq.*
7: /EMC_Celerra_SIDS3/ptocdata/1/pubpna/US11_NEW_PUB_seq.*
8: /EMC_Celerra_SIDS3/ptocdata/1/pubpna/US11_NEW_PUB_seq.*
9: /EMC_Celerra_SIDS3/ptocdata/1/pubpna/US60_NEW_PUB_seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	13.0	1000	US-11-266-748A-200928	Sequence 200928, A
2	50.2	12.6	1000	US-11-266-748A-197937	Sequence 197937, A
3	49.4	12.4	154394	US-11-266-748A-58517	Sequence 58517, A
4	49.4	12.4	394191	US-10-506-549-3	Sequence 3, Appl1
5	49.2	12.3	755217	US-11-266-748A-29045	Sequence 29045, A
6	48.6	12.2	163354	US-11-266-748A-23437	Sequence 23437, A
7	48.4	12.1	1000	US-11-266-748A-223388	Sequence 223388, A
8	48.4	12.1	1000	US-11-266-748A-290883	Sequence 290883, A
9	48.4	12.1	1000	US-11-266-748A-342312	Sequence 342312, A
10	48.4	12.1	1000	US-11-266-748A-402247	Sequence 402247, A
11	48.4	12.1	1000	US-11-266-748A-473393	Sequence 473393, A
12	48.4	12.0	2950	US-11-266-748A-24706	Sequence 24706, A
13	47.4	11.9	1000	US-11-266-748A-203031	Sequence 203031, A
14	47.2	11.8	1000	US-11-266-748A-282791	Sequence 282791, A
15	47.2	11.8	1000	US-11-266-748A-309431	Sequence 309431, A
16	47.2	11.8	1000	US-11-266-748A-309431	Sequence 309431, A
17	46.8	11.7	1000	US-11-266-748A-205479	Sequence 205479, A
18	46.8	11.7	376	US-11-266-748A-391612	Sequence 391612, A
19	46.6	11.7	376	US-11-266-748A-482330	Sequence 482330, A
20	46.6	11.7	1000	US-11-266-748A-281212	Sequence 281212, A
21	46.6	11.7	1000	US-11-266-748A-307852	Sequence 307852, A
22	46.6	11.7	1000	US-11-266-748A-390020	Sequence 390020, A
23	46.6	11.7	1000	US-11-266-748A-480738	Sequence 480738, A
24	46.6	11.7	3564	US-11-266-748A-25970	Sequence 25970, A

25	46.2	11.6	3537	US-11-266-748A-29487	Sequence 29487, A
26	46.2	11.6	21466	US-11-266-748A-22658	Sequence 22658, A
27	46	11.5	12423	US-10-517-441-251	Sequence 251, App
28	46	11.5	12423	US-10-517-441-251	Sequence 251, App
29	46	11.5	118384	US-11-266-748A-23687	Sequence 23687, A
30	46	11.5	634888	US-10-533-365-1	Sequence 1, Appl1
31	45.8	11.5	1000	US-11-266-748A-293635	Sequence 293635, A
32	45.8	11.5	1000	US-11-266-748A-345064	Sequence 345064, A
33	45.8	11.5	1000	US-11-266-748A-405114	Sequence 405114, A
34	45.8	11.5	1000	US-11-266-748A-476760	Sequence 476760, A
35	45.8	11.5	5823	US-11-266-748A-31555	Sequence 31555, A
36	45.8	11.5	5823	US-11-266-748A-57556	Sequence 57556, A
37	45.8	11.5	6453	US-11-266-748A-31554	Sequence 31554, A
38	45.6	11.4	1000	US-11-266-748A-117839	Sequence 117839, A
39	45.6	11.4	1000	US-11-266-748A-160003	Sequence 160003, A
40	45.6	11.4	1000	US-11-266-748A-287071	Sequence 287071, A
41	45.6	11.4	1000	US-11-266-748A-290324	Sequence 290324, A
42	45.6	11.4	1000	US-11-266-748A-338500	Sequence 338500, A
43	45.6	11.4	1000	US-11-266-748A-341753	Sequence 341753, A
44	45.6	11.4	1000	US-11-266-748A-397664	Sequence 397664, A
45	45.6	11.4	1000	US-11-266-748A-401556	Sequence 401556, A
46	45.6	11.4	1000	US-11-266-748A-468710	Sequence 468710, A
47	45.6	11.4	1000	US-11-266-748A-472602	Sequence 472602, A
48	45.6	11.4	201144	US-11-266-748A-23494	Sequence 23494, A
49	45.4	11.4	740	US-11-266-748A-53653	Sequence 53653, A
50	45.4	11.4	2037	US-11-266-748A-31003	Sequence 31003, A
51	45.2	11.3	1421559	US-11-266-748A-228208	Sequence 228208, A
52	45.2	11.3	38945	US-11-266-748A-24969	Sequence 24969, A
53	44.8	11.2	502	US-11-266-748A-198276	Sequence 198276, A
54	44.8	11.2	1000	US-11-266-748A-23990	Sequence 23990, A
55	44.8	11.2	42790	US-11-266-748A-24242	Sequence 24242, A
56	44.8	11.2	178403	US-11-266-748A-22732	Sequence 22732, A
57	44.6	11.2	14983	US-11-347-766-77	Sequence 77, Appl1
58	44.6	11.2	84248	US-11-266-748A-197513	Sequence 197513, A
59	44.4	11.1	1000	US-11-266-748A-23474	Sequence 23474, A
60	44.4	11.1	347503	US-11-266-748A-195590	Sequence 195590, A
61	44.2	11.1	1000	US-11-266-748A-203426	Sequence 203426, A
62	44.2	11.1	96217	US-11-266-748A-59101	Sequence 59101, A
63	44.2	11.1	139947	US-11-266-748A-24987	Sequence 24987, A
64	44.2	11.1	170869	US-11-266-748A-25150	Sequence 25150, A
65	44.2	11.1	1000	US-11-266-748A-199669	Sequence 199669, A
66	44	11.0	1000	US-11-266-748A-281845	Sequence 281845, A
67	44	11.0	1000	US-11-266-748A-308485	Sequence 308485, A
68	44	11.0	1000	US-11-266-748A-27525	Sequence 27525, A
69	44	11.0	3106	US-11-266-748A-27825	Sequence 27825, A
70	44	11.0	4639	US-11-266-748A-24847	Sequence 24847, A
71	44	11.0	26757	US-11-266-748A-60803	Sequence 60803, A
72	44	11.0	176928	US-11-266-748A-43552	Sequence 43552, A
73	43.8	11.0	809	US-11-266-748A-287824	Sequence 287824, A
74	43.8	11.0	1000	US-11-266-748A-339253	Sequence 339253, A
75	43.8	11.0	1000	US-11-266-748A-396596	Sequence 396596, A
76	43.8	11.0	1000	US-11-266-748A-469642	Sequence 469642, A
77	43.8	11.0	1000	US-11-266-748A-23094	Sequence 23094, A
78	43.8	11.0	4560	US-11-266-748A-30273	Sequence 30273, A
79	43.8	11.0	5541	US-11-266-748A-56385	Sequence 56385, A
80	43.8	11.0	5541	US-11-266-748A-56385	Sequence 56385, A
81	43.8	11.0	15786	US-11-222-810-16	Sequence 16, Appl1
82	43.6	10.9	141	US-11-266-748A-11970	Sequence 11970, A
83	43.6	10.9	1000	US-11-266-748A-12001	Sequence 12001, A
84	43.6	10.9	1000	US-11-266-748A-200361	Sequence 200361, A
85	43.6	10.9	1000	US-11-266-748A-282919	Sequence 282919, A
86	43.6	10.9	1000	US-11-266-748A-309559	Sequence 309559, A
87	43.6	10.9	1000	US-11-266-748A-392519	Sequence 392519, A
88	43.6	10.9	1000	US-11-266-748A-482327	Sequence 482327, A
89	43.6	10.9	1000	US-11-266-748A-26100	Sequence 26100, A
90	43.6	10.9	3881	US-10-505-928-607	Sequence 607, App
91	43.6	10.9	144649	US-11-266-748A-59314	Sequence 59314, A
92	43.6	10.9	1000	US-11-266-748A-281299	Sequence 281299, A
93	43.4	10.9	1000	US-11-266-748A-307939	Sequence 307939, A
94	43.4	10.9	1000	US-11-266-748A-390154	Sequence 390154, A
95	43.4	10.9	1000	US-11-266-748A-480872	Sequence 480872, A
96	43.4	10.9	1000	US-11-266-748A-480872	Sequence 480872, A
97	43.4	10.9	3407	US-11-266-748A-24366	Sequence 24366, A

```

PRIORITY APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662,276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700,293
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 48396
SOFTWARE: PatentIn version 3.3
SEQ ID NO 200928
LENGTH: 1000
TYPE: DNA
ORGANISM: Homo Sapiens
US-11-266-748A-200928

Query Match      13.0%; Score 52; DB 8; Length 1000;
Best Local Similarity    55.0%; Pred. No. 5.6e-05;
Matches 122; Conservative 1; Mismatches 96; Indels 3; Gaps 1

QY      121 CCTACCTTTTGTACTGTTTGATATATATATATATACATACACACACACACACACA 180
DB      787 CTTCTTTTAAATATATATATACATATACATACACACACACACACACACACACACA 728
QY      181 CAVATATTATTAAGTCGGACATATAATGTACACCACCTATTCCAAAGTAATACTACTATCT 240
DB      727 CACATATATATAAATATACACATATATATGTACAC---TTCTATATAAAAATTTATGTT 671
QY      241 CAGAAGTAAGACATATATCTGCTATTTGTCAATTAATCAGAGTTAAGCTGCCTAACCTTT 300
DB      670 CACATGTACAAATTTAAACAGATGTTTTTTTAAAGCAAATTAATAATATTTTAAATTT 611
QY      301 ATGCAGTTCAAGTTTCCCCTCATATGAAAAACCGAAGGTTGAA 342
DB      610 CTCAGTTTCAGTTTCCAAATATGATTAATATTTGATGATATATA 569

RESULT 2
US-11-266-748A-197937
Sequence 197937, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcriptional Microarray Technology and Methods of Using the Same
```



```
TYPE: DNA
ORGANISM: Homo Sapiens
US-11-266-748A-23045

Query Match
Best Local Similarity 12.3%; Score 49.2; DB 8; Length 755217;
Matches 87; Conservative 1; Mismatches 64; Indels 0; Gaps 0;

98 GTAGGTGACGAGGAGTGGAGGCTCTTTTGTACTGTTGATATATATATACAT 157
DB GCATTATATATGATGCACTTCATGCAATATGATGCTGATATATATATATAC 87212

158 ACACACACACACACACACACACACACATATTTTACTGCGACATAATGACACACCT 217
DB ACACACACACACACACACACACACACACACACACACACACACACACACACAC 87211

218 ATTCAAGTAAATCTACTATCTCAGAACTTA 249
DB AATTAATGCTACATTTTGTTTTAACTTA 87120

RESULT 6
US-11-266-748A-23437
Sequence 23437, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
TITLE OF INVENTION: Transcriptome Microarray Technology and
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
CURRENT FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662,276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700,293
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 483996
SOFTWARE: PatentIn version 3.3
SEQ ID NO 23437
LENGTH: 162354
TYPE: DNA
ORGANISM: Homo Sapiens
US-11-266-748A-23437

Query Match
Best Local Similarity 12.2%; Score 48.6; DB 8; Length 162354;
Matches 66; Conservative 1; Mismatches 30; Indels 0; Gaps 0;
```

```
Sequence 223388, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcriptome Microarray Technology and
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
CURRENT FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662,276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700,293
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 483996
SOFTWARE: PatentIn version 3.3
SEQ ID NO 223388
LENGTH: 1000
TYPE: DNA
ORGANISM: Homo Sapiens
US-11-266-748A-223388

Query Match
Best Local Similarity 12.1%; Score 48.4; DB 8; Length 1000;
Matches 52; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

130 TTGTACTGTTGATATATATATATATACATACACACACACACACACATATTT 189
DB TTGTTTCATCATATATATATATATACACACACACACACACACACATATTTT 585

RESULT 8
US-11-266-748A-290883
Sequence 290883, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcriptome Microarray Technology and
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
CURRENT FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662,276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700,293
PRIOR FILING DATE: 2005-07-18
```

NUMBER OF SEQ ID NOS: 483996
SOFTWARE: PatentIn version 3.3
SEQ ID NO 290883
LENGTH: 1000
TYPE: DNA
ORGANISM: Homo Sapiens
US-11-266-748A-290883

Query Match 12.1% Score 48.4; DB 8; Length 1000;
Best Local Similarity 86.7%; Pred. No. 0.00058;
Matches 52; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 130 TTGTTCTGTTGATATATATATATATACACACACACACACACACATATTTT 189
DB 526 TTGTTCTGATCATATATATATATATACACACACACACACACACATATTTT 585

RESULT 9

US-11-266-748A-342312/c
Sequence 342312, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcriptome Microarray Technology and
TITLE OF INVENTION: Methods of Using the Same
FILE REFERENCE: 55815-0102 (319189)
CURRENT FILING DATE: 2005-11-03
PRIOR FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662,276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700,293
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 483996
SOFTWARE: PatentIn version 3.3
SEQ ID NO 342312
LENGTH: 1000
TYPE: DNA
ORGANISM: Homo Sapiens
US-11-266-748A-342312

Query Match 12.1% Score 48.4; DB 8; Length 1000;
Best Local Similarity 86.7%; Pred. No. 0.00058;
Matches 52; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 130 TTGTTCTGTTGATATATATATATATACACACACACACACACACATATTTT 189
DB 475 TTGTTCTGATCATATATATATATATACACACACACACACACACATATTTT 416

RESULT 10

US-11-266-748A-402247
Sequence 402247, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcriptome Microarray Technology and
TITLE OF INVENTION: Methods of Using the Same

US-11-266-748A-402247
Sequence 402247, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcriptome Microarray Technology and
TITLE OF INVENTION: Methods of Using the Same

FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
CURRENT FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662,276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700,293
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 483996
SOFTWARE: PatentIn version 3.3
SEQ ID NO 402247
LENGTH: 1000
TYPE: DNA
ORGANISM: Homo Sapiens
US-11-266-748A-402247

Query Match 12.1% Score 48.4; DB 8; Length 1000;
Best Local Similarity 86.7%; Pred. No. 0.00058;
Matches 52; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 130 TTGTTCTGTTGATATATATATATATACACACACACACACACACATATTTT 189
DB 526 TTGTTCTGATCATATATATATATATACACACACACACACACACATATTTT 585

RESULT 11

US-11-266-748A-473293/c
Sequence 473293, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcriptome Microarray Technology and
TITLE OF INVENTION: Methods of Using the Same
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
CURRENT FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662,276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700,293
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 483996
SOFTWARE: PatentIn version 3.3
SEQ ID NO 473293
LENGTH: 1000
TYPE: DNA
ORGANISM: Homo Sapiens
US-11-266-748A-473293

US-11-266-748A-473293
Sequence 473293, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcriptome Microarray Technology and
TITLE OF INVENTION: Methods of Using the Same

	Query Match	12.1%	Score 48.4	DB 8	Length 1000;
	Best Local Similarity	86.7%	Pred. No. 0.00058		
	Matches	52;	Conservative	1; Mismatches 7; Indels 0; Gaps 0;	
Oy		130	TTTGACGTTGGATATATATATATATATACACACACACACACACATATT		189
Dp		475	TTGTTCATCATATATATATATATATACACACACACACACACACATATT		416

RESULT 12
US-11-266-748A-24706
Sequence 24706, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
TITLE OF INVENTION: Mulligan, Karl
TITLE OF INVENTION: Transcibptome Microarray Technology and
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266, 748A
CURRENT FILING DATE: 2005-11-03

OY	166	ACACACACACACACAYATAT	187
		:	
Dd	2485	ACACACACACACACATATAT	2506

RESULT 13
US-11-266-748A-203031/C
Sequence 203031, Application US/11266748A
Publication No. US2006013463A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcription Microarray Technology and
TITLE OF INVENTION: Methods of Using the Same
FILE REFERENCE: 5815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
CURRENT FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2

```

1      PRIOR FILING DATE: 2004-11-03
2      PRIOR APPLICATION NUMBER: EP 04105482.6
3      PRIOR FILING DATE: 2004-11-03
4      PRIOR APPLICATION NUMBER: EP 04105483.4
5      PRIOR FILING DATE: 2004-11-03
6      PRIOR APPLICATION NUMBER: EP 04105507.0
7      PRIOR FILING DATE: 2004-11-03
8      PRIOR APPLICATION NUMBER: EP 04105485.9
9      PRIOR FILING DATE: 2004-11-03
10     PRIOR APPLICATION NUMBER: EP 04105484.2
11     PRIOR FILING DATE: 2004-11-03
12     PRIOR APPLICATION NUMBER: US 60/662,276
13     PRIOR FILING DATE: 2005-03-14
14     PRIOR APPLICATION NUMBER: US 60/700,293
15     PRIOR FILING DATE: 2005-07-18
16     NUMBER OF SEQ ID NOS: 483936
17     SOFTWARE: Patentin version 3.3
18     SEQ ID NO 203031
19     LENGTH: 1000
20     TYPE: DNA
21     ORGANISM: Homo Sapiens
22     US-11-266-748A-203031

```

Query Match	11.93	Score 47.4	DB 8	Length 1000
Best Local Similarity	62.63	Pred. No. 0.0011		
Matches 72; Conservative	1	Mismatches 42	Indels 0	Gaps 0

QY 183 YATATTTAACTCGGACATAAATGTACACACCTATTCAAGTAAAACTACTA 237
:
670 CACACACTCAGTATTGAAAAACAATGAAAAACATTTAAAAATACACATACCTTTTA 616
Db

RESULT 14
US-11-266-748A-282791
Sequence 282791, Application US/11266748A
Publication No. US2006013465A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcription Microarray Technology and
TITLE OF INVENTION: Methods of Using the Same


```

? APPLICANT: Harkin, Paul
? APPLICANT: Johnston, Patrick
? APPLICANT: Mulligan, Karl
? TITLE OF INVENTION: Transcriptome Microarray Technology and
? TITLE OF INVENTION: Methods of Using the Same
? FILE REFERENCE: 55815-0102 (319189)
? CURRENT APPLICATION NUMBER: US/11/266,748A
? CURRENT FILING DATE: 2005-11-03
? PRIOR APPLICATION NUMBER: EP 04105479.2
? PRIOR FILING DATE: 2004-11-03
? PRIOR APPLICATION NUMBER: EP 04105482.6
? PRIOR FILING DATE: 2004-11-03
? PRIOR APPLICATION NUMBER: EP 04105483.4
? PRIOR FILING DATE: 2004-11-03
? PRIOR APPLICATION NUMBER: EP 04105507.0
? PRIOR FILING DATE: 2004-11-03
? PRIOR APPLICATION NUMBER: EP 04105485.9
? PRIOR FILING DATE: 2004-11-03
? PRIOR APPLICATION NUMBER: EP 04105484.2
? PRIOR FILING DATE: 2004-11-03
? PRIOR APPLICATION NUMBER: US 60/662,276
? PRIOR FILING DATE: 2005-03-14
? PRIOR APPLICATION NUMBER: US 60/700,293
? PRIOR FILING DATE: 2005-07-18
? NUMBER OF SEQ ID NOS: 483596
? SOFTWARE: PatentIn version 3.3
? SEQ ID NO 281212
? LENGTH: 1000
? TYPE: DNA
? ORGANISM: Homo Sapiens
? US-11-266-748A-281212

```

Query Match	Similarity	11.7%	Score 46.6	DB 8	Length 1000
Best Local	Similarity	70.1%	Pred. No. 0.0018		
Matches	61	Conservative	1	Mismatches	25
				Indels	0
				Gaps	0
QY	142 ATATATATATATATACATACACACACACACACACATATATTTAACTCGGACA	201			
DB	866 ATATATATACATACACACACACACATATATATACACACACACATATATATACACACA	807			
QY	202 TTAATGTACACCACTTATTCAAAGTAA	228			
DB	806 TACATACACACATGCTTTAAAAATCA	780			
RESULT 21					
US-11-266-748A-307852	Application US/11266748A				
Sequence 307852, US20060134663A1	Publication No. US20060134663A1				
GENERAL INFORMATION:					
APPLICANT: Harkin, Paul					
APPLICANT: Johnston, Patrick					
APPLICANT: Mulligan, Karl					
TITLE OF INVENTION: Transcriptome Microarray Technology and					
FILE OF INVENTION: Methods of Using the Same					
FILE REFERENCE: 55815-0102 (319189)					
CURRENT APPLICATION NUMBER: US/11/266, 748A					
CURRENT FILING DATE: 2005-11-03					
PRIOR APPLICATION NUMBER: EP 04105479.2					
PRIOR FILING DATE: 2004-11-03					
PRIOR APPLICATION NUMBER: EP 04105482.6					
PRIOR FILING DATE: 2004-11-03					
PRIOR APPLICATION NUMBER: EP 04105483.4					
PRIOR FILING DATE: 2004-11-03					
PRIOR APPLICATION NUMBER: EP 04105507.0					
PRIOR FILING DATE: 2004-11-03					
PRIOR APPLICATION NUMBER: EP 04105485.9					
PRIOR FILING DATE: 2004-11-03					
PRIOR APPLICATION NUMBER: EP 04105484.2					
PRIOR FILING DATE: 2004-11-03					
PRIOR APPLICATION NUMBER: US 60/662, 276					
PRIOR FILING DATE: 2005-03-14					
PRIOR APPLICATION NUMBER: US 60/700, 293					

```

; PRIOR FILING DATE: 2005-07-18
;
; NUMBER OF SEQ ID NOS: 483396
;
; SOFTWARE: PatSeqin version 3.3.3
;
; SEQ ID NO 307852
;
; LENGTH: 1000
;
; TYPE: DNA
;
; ORGANISM: Homo Sapiens
;
US-11-266-748A-307852

```

	Query Match	11.7%;	Score 46.6;	DB 8;	Length 1000;	
	Best Local Similarity	70.1%;	Pred. No. 0.0018;			
	Matches	61;	Conservative	1;	Mismatches	25; Indels 0; Gaps 0
Oy	142	ATATATTATTTACATACACACACACACACACACACACACATATTTAACTGGGACA	201			
Dd	135	ATAATTCATCATCACACACACACACATTTATTCACACACACATAGTATTTATACACA	194			
Oy	202	TAAATGTACACCACCTATTCAAAGTAA	228			
Dd	195	TACATACACACATGCCTTTAAAAATCA	221			

```

RESULT 22
US-11-266-748A-390020/c
; Sequence 390020, Application US//11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Hartlin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US//11266, 748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 350020
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-390020

Query Match 11.7%; Score 46.6; DB 8; Length 1000;
Best Local Similarity 70.1%; Pred. No. 0.0018;
Matches 61; Conservative 1; Mismatches 25; Indels 0; Gaps 0;

QY 142 ATATATATATATATCATACACACACACACACACACATATTTTAAGTGGGACA 201
DB 866 ATATATATATATATCATACACACACACACATATATATATATATATATACACACA 807
QY 202 TAAATGTACACCAACTTATTCAGAGTAA 228
DB 806 TACATACACACATGCTTTAAAAATCA 780

RESULT 23

```

```

US-11-266-748A-480738
Sequence 480738, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
TITLE OF INVENTION: Transcription Microarray Technology and
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
SEQU ID NO 480738
LENGTH: 1000
TYPE: DNA
ORGANISM: Homo Sapiens
US-11-266-748A-480738
Query Match 11.7%; Score 46.6; DB 8; Length 1000;
Beat Local Similarity 70.1%; Pred. No. 0.0018;
Matches 61; Conservative 1; Mismatches 25; Indels 0; Gaps 0;
QY 142 ATATATATATATACATACACACACACACACACATATTTAACTCGGAC 201
DB 135 ATATATACATACACACACACACATATATACACACACATAGTATATACAC 194
QY 202 TAAATGTACACCACCTATTCAAGTAA 228
DB 195 TACATACACACATGCGCTTTAAATAACA 221
RESULT 24
US-11-266-748A-25970/c
Sequence 25970, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcription Microarray Technology and
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2

```

```

; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25970
; LENGTH: 3564
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-25970

Query Match      11.7%; Score 46.6; DB 8; Length 3564;
Best Local Similarity 65.0%; Pred. No. 0.003;
Matches 67; Conservative 1; Mismatches 35; Indels 0; Gaps 0;

OY    152 ATACATCACACACACACACACAAGTATTCTGCGGACATAATGTACA 211
DB    2349 ACACACACACACACACACACACACACACACACACACACAAATGAAGAAGACAGAGACA 2280
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY    212 CCACCTATTCAAAGTAATAACTACTATTCGAGAAGTAAGACA 254
DB    2289 TATATAAACGATATAAAAAGTAGGGCCCAAGATGTACAAAAA 2247

RESULT 25
US-11-266-748A-29487/c
Sequence 29487, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcriptome Microarray Technology and
FILE REFERENCE: 55815-0102 (319189)
CURRENT FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479. 2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482. 6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483. 4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507. 0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485. 9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484. 2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662,276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700,293
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 483996
SOFTWARE: PatentIn version 3.3
SEQ ID NO 29487
LENGTH: 3537
TYPE: DNA
ORGANISM: Homo Sapiens
US-11-266-748A-29487

Query Match      11.6%; Score 46.2; DB 8; Length 3537;
Best Local Similarity 65.3%; Pred. No. 0.0039;
Matches 66; Conservative 1; Mismatches 34; Indels 0; Gaps 0;

OY    154 ACATACACACACACACACACACACACACAVTATTCTGCGGACATAATGTACACC 213
DB    2347 ACAACACACACACACACACACACACACACACACACAAATGAAGAAGACAGAGACATA 2288
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY    214 AACTTTCAAAGTAAAAAGTGATCTATCTCAGAAAGTAAAGACA 254
```



```
; APPLICANT: MARY, Almuth
; APPLICANT: HOEFLE, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
; TITLE OF INVENTION: Proliferative disorders
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; PRIOR FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; SEQUENCE ID NO 251
; LENGTH: 12423
; TYPE: DNA
; ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-251

Query Match      11.5%; Score 46; DB 6; Length 12423;
Best Local Similarity 59.4%; Pred. No. 0.0071;
Matches 76; Conservative 1; Mismatches 51; Indels 0; Gaps 0;

OY       139 TTGAATATTATTTATTCATAACAGCACACACACACACACACACACAYATTTTAACTCGGG    198
          |||||
DB        7720 TTACAACACACACACACACACACACACACACACACACACACACACTTCCG    7661

OY       199 ACATAAATGTACACCACCACTATTCAAAGTAAGAAAAGTAAAGCATTAAT    258
          |||||||
DB        7660 ACTTAATATTAATCTCCTCTCTCTAAAAAACCTTATCCTTAAAAAAAAAAAAACAAAA    7601

OY       259 CCTGCCTAT   266
          |         |||
DB        7600 CAAATAT   7593

RESULT 28
US-10-517-441-525/c
; Sequence 525, Application US/10517441
; Publication No.: US20060121467A1
GENERAL INFORMATION:
APPLICANT: FOEKENS, John
APPLICANT: HAREBECK, Nadia
APPLICANT: KOENIG, Thomas
APPLICANT: MAIER, Sabine
APPLICANT: MARTENS, John
APPLICANT: MODEL, Fabian
APPLICANT: NIMWICH, Inko
APPLICANT: RUJAN, Tamas
APPLICANT: SCHMITZ, Armin
APPLICANT: SCHMIDT, Manfred
APPLICANT: LOOK, Maxime P.
APPLICANT: MARK, Almuth
APPLICANT: HOEFLER, Heinz
TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
FILE REFERENCE: 47675-93
CURRENT APPLICATION NUMBER: US/10/517,441
PRIOR FILING DATE: 2004-12-11
PRIOR APPLICATION NUMBER: PCT/EP2003/010881
PRIOR FILING DATE: 2003-10-01
PRIOR APPLICATION NUMBER: DE 10317955.0
PRIOR FILING DATE: 2003-04-17
PRIOR APPLICATION NUMBER: DE 10300096.8
PRIOR FILING DATE: 2003-01-07
PRIOR APPLICATION NUMBER: DE 10245779.4
PRIOR FILING DATE: 2002-10-01
SEQUENCE ID NOS: 2147
```

```
/ SEQ ID NO 525
/ LENGTH: 12423
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-525
```

```
Query Match      11.5%; Score 46; DB 6; Length 12423;
Best Local Similarity 59.4%; Pred. No. 0.0071;
Matches 76; Conservative 1; Mismatches 51; Indels 0; Gaps 0;
```

```
QY 139 TTGATATATATATATATACATACACACACACACACACATATTTTAACTCGGG 198
DB 7720 TTACACACACACACACACACACACACACACACACACACACACACACCTTCC 7661
QY 199 ACCTAATATGACACACACCTTTTCAAAAGTAAACTACTATCTCGAAGTAAAGCATAT 258
DB 7660 ACCTAATCTAAATCTCTCTCTTAAAAAACCTATCCCTTAAAAAACCACAAAA 7601
QY 259 CCTGCTAT 266
DB 7600 CAAATAT 7593
```

RESULT 29

```
US-11-266-748A-23687/C
/ Sequence 23687, Application US/11266748A
/ Publication No. US20060134663A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Harkin, Paul
/ APPLICANT: Johnston, Patrick
/ APPLICANT: Mulligan, Karl
/ TITLE OF INVENTION: Transcription Microarray Technology and
/ FILE REFERENCE: 55815-0102 (319189)
/ CURRENT APPLICATION NUMBER: US/11/266,748A
/ PRIOR FILING DATE: 2005-11-03
/ PRIOR APPLICATION NUMBER: EP 04105479.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105482.6
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105483.4
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105507.0
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105485.9
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105484.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: US 60/662,276
/ PRIOR FILING DATE: 2005-03-14
/ PRIOR APPLICATION NUMBER: US 60/700,293
/ PRIOR FILING DATE: 2005-07-18
/ NUMBER OF SEQ ID NOS: 48396
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 23687
/ LENGTH: 118384
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-11-266-748A-23687
```

```
Query Match      11.5%; Score 46; DB 8; Length 118384;
Best Local Similarity 64.4%; Pred. No. 0.017;
Matches 67; Conservative 1; Mismatches 36; Indels 0; Gaps 0;
```

```
QY 153 TACATACACACACACACACACACACACACATATTTTAACTCGGACATTAATGTAC 212
DB 115758 TACACACACACACACACACACACACACACATTTATCAATGTATATACAGTCAT 115699
QY 213 CACCTATTTCAAGTAAACTACTATCTCGAAGTAAAGCATATA 256
DB 115698 GTGTCAATTCATGACAAAGATCTTTCTGAGAAATGTATCTCTTA 115655
```

RESULT 30

```
US-10-533-365-1/C
/ Sequence 1, Application US/10533365
/ Publication No. US20060141462A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Reynisdottir, Inga
/ APPLICANT: Gulcher, Jeffrey R.
/ APPLICANT: Grant, Struan F.
/ APPLICANT: Thorleifsson, Gudmar
/ TITLE OF INVENTION: Human Type II Diabetes Gene - SLC-3
/ FILE REFERENCE: 2345, 2046-007
/ CURRENT APPLICATION NUMBER: US/10/533,365
/ CURRENT FILING DATE: 2005-04-29
/ PRIOR APPLICATION NUMBER: PCT/US03/34801
/ PRIOR FILING DATE: 2003-10-31
/ PRIOR APPLICATION NUMBER: 60/423,541
/ PRIOR FILING DATE: 2002-11-01
/ NUMBER OF SEQ ID NOS: 236
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 634888
/ TYPE: DNA
/ ORGANISM: homo sapiens
US-10-533-365-1
```

```
Query Match      11.5%; Score 46; DB 6; Length 634888;
Best Local Similarity 62.5%; Pred. No. 0.032;
Matches 70; Conservative 1; Mismatches 41; Indels 0; Gaps 0;
```

```
QY 95 AAAGTAGTATGACGAGGAGGAGGCTTTTGTACTGTGATATATATATATATATAT 154
DB 102116 AAAGTAGTATGACGAGGAGGAGGCTTTTGTACTGTGATATATATATATATATAT 102057
QY 155 CATACACACACACACACACACACACACATATTTTAACTCGGACATTAAT 206
DB 102056 AATACACACACACACACACACATATATATATATATATATATATATATATATATAT 102005
```

RESULT 31

```
US-11-266-748A-293635/C
/ Sequence 293635, Application US/11266748A
/ Publication No. US20060134663A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Harkin, Paul
/ APPLICANT: Johnston, Patrick
/ APPLICANT: Mulligan, Karl
/ TITLE OF INVENTION: Transcription Microarray Technology and
/ FILE REFERENCE: 55815-0102 (319189)
/ CURRENT APPLICATION NUMBER: US/11/266,748A
/ PRIOR FILING DATE: 2005-11-03
/ PRIOR APPLICATION NUMBER: EP 04105479.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105482.6
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105483.4
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105507.0
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105485.9
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105484.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: US 60/662,276
/ PRIOR FILING DATE: 2005-03-14
/ PRIOR APPLICATION NUMBER: US 60/700,293
/ PRIOR FILING DATE: 2005-07-18
/ NUMBER OF SEQ ID NOS: 48396
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 293635
```

[illegible]

```

      / APPLICANT: Hartkin, Paul
      / APPLICANT: Johnston, Patrick
      / APPLICANT: Mulligan, Karl
      / TITLE OF INVENTION: Transcriptome Microarray Technology and
      / TITLE OF INVENTION: Methods of Using the Same
      / FILE REFERENCE: 55815-0102 (319189)
      / CURRENT APPLICATION NUMBER: US/11/666,748A
      / CURRENT FILING DATE: 2005-11-03
      / PRIOR APPLICATION NUMBER: EP 04105479.2
      / PRIOR FILING DATE: 2004-11-03
      / PRIOR APPLICATION NUMBER: EP 04105482.6
      / PRIOR FILING DATE: 2004-11-03
      / PRIOR APPLICATION NUMBER: EP 04105485.9
      / PRIOR FILING DATE: 2004-11-03
      / PRIOR APPLICATION NUMBER: EP 04105507.0
      / PRIOR FILING DATE: 2004-11-03
      / PRIOR APPLICATION NUMBER: EP 04105483.4
      / PRIOR FILING DATE: 2004-11-03
      / PRIOR APPLICATION NUMBER: EP 04105507.0
      / PRIOR FILING DATE: 2004-11-03
      / PRIOR APPLICATION NUMBER: EP 04105485.9
      / PRIOR FILING DATE: 2004-11-03
      / PRIOR APPLICATION NUMBER: EP 04105484.2
      / PRIOR FILING DATE: 2004-11-03
      / PRIOR APPLICATION NUMBER: US 60/700,293
      / PRIOR FILING DATE: 2005-07-18
      / NUMBER OF SEQ ID NOS: 483996
      / SOFTWARE: PatentIn version 3.3
      / SEQ ID NO 405714
      / LENGTH: 1000
      / TYPE: DNA
      / ORGANISM: Homo Sapiens
US-11-266-748A-405714

Query Match          11.5%; Score 45.8; DB 8; Length 1000;
Beet Local Similarity 79.1%; Pred.No. 0.0031;
Matches 53; Conservative 1; Mismatches 13; Indels 0; Gaps 0

QY      123 TACCTTTTGACTGTTGGATATATATATATACATCACACACACACACACA 182
        ||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db       591 TATATAATATAATATATATATATATATATATATATATATATATATACACACACACACACACA 532

QY      183 YATATTTT 189
        :||| |
Db       531 CATATAT 525

RESULT 34
US-11-266-748A-476760
/ Sequence 476760, Application US/11266748A
/ Publication No. US2006013463A1
/ GENERAL INFORMATION:
/ APPLICANT: Hartkin, Paul
/ APPLICANT: Johnston, Patrick
/ APPLICANT: Mulligan, Karl
/ TITLE OF INVENTION: Transcriptome Microarray Technology and
/ TITLE OF INVENTION: Methods of Using the Same
/ FILE REFERENCE: 55815-0102 (319189)
/ CURRENT APPLICATION NUMBER: US/11/266,748A
/ CURRENT FILING DATE: 2005-11-03
/ PRIOR APPLICATION NUMBER: EP 04105479.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105482.6
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105483.4
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105507.0
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105485.9
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105484.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: US 60/662,276
/ PRIOR FILING DATE: 2005-03-14
/ PRIOR APPLICATION NUMBER: US 60/700,293

```


RESULT 39

;
; PRIOR APPLICATION NUMBER: US 60/662,270
;
; PRIOR FILING DATE: 2005-03-14
;
; PRIOR APPLICATION NUMBER: US 60/700,293

```

; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 287071
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-287071

Query Match      11.4% Score 45.6; DB 8; Length 1000;
Best Local Similarity 77.1%; Pred. No. 0.0035;
Matches 54; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

QY 136 TGGTTGATATATATATACACACACACACACACACATATTTTAACTC 195
Db 299 TGAATTTAAAAATATATATATATACACACACACACACACATTTTATC 240

QY 196 GGGACATATA 205
Db 239 AGGAGAGAAA 230

RESULT 41
US-11-266-748A-290324
; Sequence 290324, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 290324
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-290324

Query Match      11.4% Score 45.6; DB 8; Length 1000;
Best Local Similarity 88.9%; Pred. No. 0.0035;
Matches 48; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 136 TGGTTGATATATATATACACACACACACACACACATATTTT 189
Db 577 TATATATATATATATATATATATACACACACACACACACATATAT 630

RESULT 42
US-11-266-748A-338500
; Sequence 338500, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 338500
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-338500

Query Match      11.4% Score 45.6; DB 8; Length 1000;
Best Local Similarity 77.1%; Pred. No. 0.0035;
Matches 54; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

QY 136 TGGTTGATATATATATATACACACACACACACACACATATTTTAACTC 195
Db 702 TGAATTTAAAAATATATATATATATACACACACACACACATTTTATC 761

QY 196 GGGACATATA 205
Db 762 AGGAGAGAAA 771

RESULT 43
US-11-266-748A-341753/C
; Sequence 341753, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
```

```
;; PRIOR FILING DATE: 2005-07-18
;; NUMBER OF SEQ ID NOS: 483996
;; SOFTWARE: Patentin version 3.3
;; SEQ ID NO 341753
;; LENGTH: 1000
;; TYPE: DNA
;; ORGANISM: Homo Sapiens
US-11-266-748A-341753

Query Match
Best Local Similarity 11.4%; Score 45.6; DB 8; Length 1000;
Matches 48; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 136 TGTGATATATATATATACACACACACACACACACATATTT 189
DB 424 TATATATATATATATATATATACACACACACACACACACATATAT 371

RESULT 44
US-11-266-748A-397664/C
;; Sequence 397664, Application US/11266748A
;; Publication No. US2006013463A1
;; GENERAL INFORMATION:
;; APPLICANT: Hartlin, Paul
;; APPLICANT: Johnston, Patrick
;; APPLICANT: Mulligan, Karl
;; TITLE OF INVENTION: Transcription Microarray Technology and
;; TITLE OF INVENTION: Methods of Using the Same
;; FILE REFERENCE: 55815-0102 (319189)
;; CURRENT APPLICATION NUMBER: US/11/266,748A
;; PRIOR FILING DATE: 2005-11-03
;; PRIOR APPLICATION NUMBER: EP 04105479.2
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 04105482.6
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 04105483.4
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 04105507.0
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 04105485.9
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 04105484.2
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 60/662,276
;; PRIOR FILING DATE: 2005-03-14
;; PRIOR APPLICATION NUMBER: US 60/700,293
;; PRIOR FILING DATE: 2005-07-18
;; NUMBER OF SEQ ID NOS: 483996
;; SOFTWARE: Patentin version 3.3
;; SEQ ID NO 397664
;; LENGTH: 1000
;; TYPE: DNA
;; ORGANISM: Homo Sapiens
US-11-266-748A-397664

Query Match
Best Local Similarity 11.4%; Score 45.6; DB 8; Length 1000;
Matches 54; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

QY 136 TGTGATATATATATATACACACACACACACACACATATTT 195
DB 299 TGAATTTAAAAATATATATATATACACACACACACACACATTTTATC 240

QY 196 GGGACATATA 205
DB 239 AGGAAAGAAA 230

RESULT 45
US-11-266-748A-401556
;; Sequence 401556, Application US/11266748A
;; Publication No. US2006013463A1
;; GENERAL INFORMATION:
```

```
;; APPLICANT: Hartlin, Paul
;; APPLICANT: Johnston, Patrick
;; APPLICANT: Mulligan, Karl
;; TITLE OF INVENTION: Transcription Microarray Technology and
;; TITLE OF INVENTION: Methods of Using the Same
;; FILE REFERENCE: 55815-0102 (319189)
;; CURRENT APPLICATION NUMBER: US/11/266,748A
;; PRIOR FILING DATE: 2005-11-03
;; PRIOR APPLICATION NUMBER: EP 04105479.2
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 04105482.6
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 04105483.4
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 04105507.0
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 04105485.9
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 04105484.2
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: US 60/662,276
;; PRIOR FILING DATE: 2005-03-14
;; PRIOR APPLICATION NUMBER: US 60/700,293
;; PRIOR FILING DATE: 2005-07-18
;; NUMBER OF SEQ ID NOS: 483996
;; SOFTWARE: Patentin version 3.3
;; SEQ ID NO 468710

RESULT 46
US-11-266-748A-468710
;; Sequence 468710, Application US/11266748A
;; Publication No. US2006013463A1
;; GENERAL INFORMATION:
;; APPLICANT: Hartlin, Paul
;; APPLICANT: Johnston, Patrick
;; APPLICANT: Mulligan, Karl
;; TITLE OF INVENTION: Transcription Microarray Technology and
;; TITLE OF INVENTION: Methods of Using the Same
;; FILE REFERENCE: 55815-0102 (319189)
;; CURRENT APPLICATION NUMBER: US/11/266,748A
;; PRIOR FILING DATE: 2005-11-03
;; PRIOR APPLICATION NUMBER: EP 04105479.2
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 04105482.6
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 04105483.4
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 04105507.0
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 04105485.9
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 04105484.2
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: US 60/662,276
;; PRIOR FILING DATE: 2005-03-14
;; PRIOR APPLICATION NUMBER: US 60/700,293
;; PRIOR FILING DATE: 2005-07-18
;; NUMBER OF SEQ ID NOS: 483996
;; SOFTWARE: Patentin version 3.3
;; SEQ ID NO 468710

QY 136 TGTGATATATATATATACACACACACACACACACATATTT 189
DB 577 TATATATATATATATATATATACACACACACACACACACATATAT 630

Query Match
Best Local Similarity 11.4%; Score 45.6; DB 8; Length 1000;
Matches 48; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```

; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-468710
Query Match
  11.4%; Score 45.6; DB 8; Length 1000;
Best Local Similarity 77.1%; Pred. No. 0.0035;
Matches 54; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

QY 136 TGTGTATATATATATATACATACACACACACACACATATTTTACTC 195
DB 702 TGAATTTAAAAAATATATATATATATACACACACACACATATTTTATC 761

QY 196 GGGACATATA 205
DB 762 AGGAGAAAAA 771

RESULT 47
US-11-266-748A-472602/c
; Sequence 472602, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE OF INVENTION: Method of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 472602
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-472602
Query Match
  11.4%; Score 45.6; DB 8; Length 1000;
Best Local Similarity 88.9%; Pred. No. 0.0035;
Matches 46; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 136 TGTGTATATATATATATACATACACACACACACACATATTTT 189
DB 424 TATATATATATATATATATATATACACACACACACACATATAT 371

RESULT 48
US-11-266-748A-23494/c
; Sequence 23494, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
```

```

; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23494
; LENGTH: 201144
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (32870)..(32870)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (32891)..(32893)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (32896)..(32896)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (32927)..(32927)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (33188)..(33188)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (33190)..(33191)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (80784)..(80803)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (98111)..(98130)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (99420)..(99421)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (99425)..(99426)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (99431)..(99431)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (99433)..(99433)
```



```

RESULT 50
US-11-266-748A-31003/C
? Sequence 31003, Application US/11266748A
? Publication No. US2006013463A1
? GENERAL INFORMATION:
? APPLICANT: Johnston, Paul
? APPLICANT: Hartin, Paul
? APPLICANT: Mulligan, Karl
? TITLE OF INVENTION: Transcriptome Microarray Technology and
? TITLE OF INVENTION: Methods of Using the Same
? FILE REFERENCE: 55815-0102 (319189)
? CURRENT APPLICATION NUMBER: US/11/566,748A
? PRIOR FILING DATE: 2005-11-03
? PRIOR APPLICATION NUMBER: EP 04105479.2
? PRIOR FILING DATE: 2004-11-03
? PRIOR APPLICATION NUMBER: EP 04105482.6
? PRIOR FILING DATE: 2004-11-03
? PRIOR APPLICATION NUMBER: EP 04105483.4
? PRIOR FILING DATE: 2004-11-03
? PRIOR APPLICATION NUMBER: EP 04105507.0
? PRIOR FILING DATE: 2004-11-03
? PRIOR APPLICATION NUMBER: EP 04105485.9
? PRIOR FILING DATE: 2004-11-03
? PRIOR APPLICATION NUMBER: EP 04105484.2
? PRIOR FILING DATE: 2004-11-03
? PRIOR APPLICATION NUMBER: US 60/662,276
? PRIOR FILING DATE: 2005-03-14
? PRIOR APPLICATION NUMBER: US 60/700,293
? PRIOR FILING DATE: 2005-07-18
? NUMBER OF SEQ ID NOS: 483996
? SOFTWARE: PatentIn version 3.3
? SEQ ID NO 31003
? LENGTH: 2037
? TYPE: DNA
? ORGANISM: Homo Sapiens
US-11-266-748A-31003

Query Match      11.4%; Score 45.4; DB 8; Length 2037;
Beat Local Similarity 63.8%; Pred. No. 0.0053;
Matches 67; Conservative 1; Mismatches 37; Indels 0; Gaps 0;

QY      146 ATATATATACATACACACACACACACACACACACACACATATTTTAACTCGGGACATATA 205
      |||||
DB      1921 ATACACACACACACACACACACACACACACACACACACACACACATGAATCTCATTTTTCATT 1862
      |||||

QY      206 TGTACACACACCTATTCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 250
      |||||
DB      1861 GGTCCATTAAATTAGAGAACCAATTAAACCACTACACGAAATTTAA 1817
      |||||

```

Search completed: July 18, 2006, 00:05:00
Job time : 262 secs